CUk: H02328-80-0261356

80

C.U.

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

WEI-LIANG CHOW
THE JOHNS HOPKINS UNIVERSITY

PHILIP HARTMAN
THE JOHNS HOPKINS UNIVERSITY

G. D. MOSTOW
YALE UNIVERSITY

STEPHEN SMALE
UNIVERSITY OF CALIFORNIA

WITH THE COÖPERATION OF

R. BOTT
C. CHEVALLEY
K. L. CHUNG
J.-I. IGUSA
A. M. GLEASON
L. V. HÖRMANDER
K. KODAIRA
D. C. SPENCER
A. WEIL

CALCU
CENTR
LIBRA
UNIVER

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LXXXVIII, Number 1
JANUARY, 1966

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND 21218
U. S. A.

CONTENTS

PAGE

Locally nice embeddings of manifolds. By J. P. HEMPEL and D. R. McMILLAN, JR., ... 1

An extension of results of Novikov and Browder. By C. T. C. WALL, 20

Derivations and automorphisms of complete regular local rings. By NICKOLAS HEEREMA, ... 33

G-induced flows. By L. AUSLANDER and L. W. GREEN, ... 43

On two problems of Mordell. By E. BOMBIERI and H. DAVENPORT, . 61

On exponential sums in finite fields. By ENRICO BOMBIERI, . . . 71

The structure of the symplectic group over a valuation ring. By CARL R. RIEHM, ... 106

On isolated rational singularities of surfaces. By MICHAEL ARTIN, . 129

On the classification of differentiable $So(n)$ actions on simply connected π-manifolds. By WU-YI HSIANG, ... 137

Ergodicity of flows on homogeneous spaces. By CALVIN C. MOORE, . 154

On maximal Riemann surfaces. By H. RENGGLI, ... 179

The generalized Smith conjecture. By CHARLES H. GRIFFIN, . . 187

Projective representation of algebraic linear groups of transformations. By T. KAMBAYASHI, ... 199

A proof of the convergence of the Poincaré-von Zeipel procedure in celestial mechanics. By RICHARD BARRAR, ... 206

On the graded ring of theta-constants (II). By JUN-ICHI IGUSA, . 221

1-1 continuous mappings onto E^n. By LESLIE C. GLASER, . . 237

Deformations of affine embeddings of complex analytic homogeneous spaces. By G. HOCHSCHILD and G. D. MOSTOW, . . 244

Isolated singularities of solutions of linear elliptic equations. By JAMES SERRIN and H. F. WEINBERGER, ... 258

The AMERICAN JOURNAL OF MATHEMATICS appears four times yearly.

The subscription price of the JOURNAL is $11.00 in the U. S.; $11.30 in Canada and $11.60 in other foreign countries. The price of single numbers is $3.00.

Manuscript intended for publication in the JOURNAL should be sent to Professor W. L. CHOW, The Johns Hopkins University, Baltimore, Md. 21218.

Subscriptions to the JOURNAL and all business communications should be sent to THE JOHNS HOPKINS PRESS, BALTIMORE, MARYLAND 21218.

Claims for replacement of missing numbers must be received within 3 months (6 months in the case of foreign subscribers) following publication of the number.

THE JOHNS HOPKINS PRESS supplies to the authors 100 free reprints of every article appearing in the AMERICAN JOURNAL OF MATHEMATICS. On the other hand, neither THE JOHNS HOPKINS PRESS nor the AMERICAN JOURNAL OF MATHEMATICS can accept orders for additional reprints. Authors interested in securing more than 100 reprints are advised to make arrangements directly with the printers, J. H. FURST CO., 109 MARKET PLACE, BALTIMORE, MARYLAND 21218.

The typescripts submitted can be in English, French, German or Italian and should be prepared in accordance with the instructions listed on the inside back cover of this issue.

Second-class postage paid at Baltimore, Maryland.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST CO., BALTIMORE, MARYLAND

AMERICAN
JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

WEI-LIANG CHOW
THE JOHNS HOPKINS UNIVERSITY

PHILIP HARTMAN
THE JOHNS HOPKINS UNIVERSITY

G. D. MOSTOW
YALE UNIVERSITY

STEPHEN SMALE
UNIVERSITY OF CALIFORNIA

WITH THE COÖPERATION OF

R. BOTT
C. CHEVALLEY
K. L. CHUNG
J.-I. IGUSA
A. M. GLEASON
L. V. HÖRMANDER
K. KODAIRA
D. C. SPENCER
A. WEIL

PUBLISHED UNDER THE JOINT AUSPICES OF
THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

VOLUME LXXXVIII

1966

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND 21218
U. S. A.

INDEX

PAGE

Hempel, J. P. and D. R. McMillan, Jr. Locally nice embeddings of manifolds, ... 1

McMillan, D. R., Jr. See J. P. Hempel and D. R. McMillan, Jr., page 1.

Wall, C. T. C. An extension of results of Novikov and Browder, . 20

Heerema, Nickolas. Derivations and automorphisms of complete regular local rings, ... 33

Auslander, L. and L. W. Green. *G*-induced flows, ... 43

Green, L. W. See L. Auslander and L. W. Green, page 43.

Bombieri, E. and H. Davenport. On two problems of Mordell, . . 61

Davenport, H. See E. Bombieri and H. Davenport, page 61.

Bombieri, Enrico. On exponential sums in finite fields, ... 71

Riehm, Carl R. The structure of the symplectic group over a valuation ring, ... 106

Artin, Michael. On isolated rational singularities of surfaces, . 129

Hsiang, Wu-yi. On the classification of differentiable $So(n)$ actions on simply connected π-manifolds, ... 137

Moore, Calvin C. Ergodicity of flows on homogeneous spaces, . . 154

Renggli, H. On maximal Riemann surfaces, ... 179

Griffin, Charles H. The generalized Smith conjecture, . . . 187

Kambayashi, T. Projective representation of algebraic linear groups of transformations, ... 199

Barrar, Richard. A proof of the convergence of the Poincaré-von Zeipel procedure in celestial mechanics, ... 206

Igusa, Jun-ichi. On the graded ring of theta-constants (II), . 221

Glaser, Leslie C. 1-1 continuous mappings onto E^n, ... 237

Hochschild, G. and G. D. Mostow. Deformations of affine embeddings of complex analytic homogeneous spaces, . . 244

Mostow, G. D. See G. Hochschild and G. D. Mostow, page 244.

Serrin, James and H. F. Weinberger. Isolated singularities of solutions of linear elliptic equations, ... 258

Weinberger, H. F. See James Serrin and H. F. Weinberger, page 258.

Hoyt, William L. On the Chow bunches for different projective embeddings of a complete variety, ... 273

PAGE

ADAMS, WILLIAM W. Transcendental numbers in the P-adic domain, 279

BURGESS, C. E. Pairs of 3-cells with intersecting boundaries in E^3, . 309

ROSENBLUM, MARVIN. A spectral theory for self-adjoint singular integral operators, ... 314

ACCOLA, R. D. M. Invariant domains for Kleinian groups, . . . 329

GERSTEN, STEPHEN M. A product formula for Wall's obstruction, . 337

FEARNLEY, L. Embeddings of topological products of K-cell-like continua, ... 347

LEVINE, HAROLD I. Mappings of manifolds into the plane, . . . 357

GRIFFITHS, PHILLIP A. The extension problem in complex analysis II; embeddings with positive normal bundle, ... 366

FALB, PETER A. A note on coverings over imperfect ground fields, . 447

ANDREWS, GEORGE E. On the theorems of Watson and Dragonette for Ramanujan' mock theta functions, ... 454

SMALE, S. Structurally stable systems are not dense, ... 491

HAMMOND, WILLIAM F. The modular groups of Hilbert and Siegel, . 497

SCHMIDT, WOLFGANG M. Simultaneous approximation to a basis of a real numberfield, ... 517

ANDRÉ, M. Categories of functors and adjoint functors, ... 529

KLEPPNER, ADAM. Representations induced from compact subgroups, 545

RIEHM, CARL R. Orthogonal groups over the integers of a local field I, 553

KADISON, RICHARD V. and JOHN R. RINGROSE. Derivations of operator group algebras, ... 562

RINGROSE, JOHN R. See Richard V. Kadison and John R. Ringrose, page 562.

HIRSCHMAN, I. I., JR. The strong Szegö limit theorem for Toeplitz determinants, ... 577

AUSLANDER, LOUIS. Modifications of solvmanifolds and G-induced flows, 615

GILBERT, JOHN E. Spectral synthesis for invariant subspaces on groups. I, ... 626

VILLANI, VINICIO. Cohomological properties of complex spaces which carry over to normalizations, ... 636

HOLMES, R. and N. SMYTHE. Algebraic invariants of isotopy of links, 646

SMYTHE, N. See R. Holmes and N. Smythe, page 646.

PORTER, GERALD J. The homotopy groups of wedges of suspensions, 655

PAGE

Hagis, Peter, Jr. Some theorems concerning partitions into odd summands, ... 664

Kodaira, K. On the structure of compact analytic surfaces, II, . 682

Beardon, A. F. The Hausdorff dimension of singular sets of properly discontinuous groups, ... 721

Sorani, Giuliano. Integral representations of holomorphic functions, 737

Artin, Michael. Lifting of two-dimensional singularities to characteristic zero, ... 747

Pollak, Barth. On the structure of local orthogonal groups, . . 763

Seeley, R. T. Singular integrals and boundary value problems, . 781

Coburn, Lewis A. The spectra of generalized Toeplitz operators, . 810

Brown, Edgar H., Jr. and Franklin P. Peterson. The Kervaire invariant of $8k+2$-manifolds, ... 815

Peterson, Franklin P. See Edgard H. Brown, Jr. and Franklin P. Peterson, page 815.

Gordon, Hugh. The maximal ideal space of a ring of measurable functions, ... 827

Andrews, George E. An analytic proof of the Rogers-Ramanujan-Gordon identities, ... 844

Hochschild, G. and G. D. Mostow. Representative algebraic structures on complex analytic homogeneous spaces, . . 847

Mostow, G. D. See G. Hochschild and G. D. Mostow, page 847.

Earle, Clifford. J. A reproducing formula for integrable automorphic forms, ... 867

Rieffel, Marc A. On extensions of locally compact groups, . . 871

Farkas, Hershel M. Special divisors and analytic subloci of Teichmueller space, ... 881

Artin, M. Etale coverings of schemes over Hensel rings, . . 915

Sakai, Schôichirô. On a problem of Calkin, ... 935

Kirwan, William E. Extremal problems for the typically real functions, ... 942

CALCUT
CENTRAL LIBRARY
UNIVER

LOCALLY NICE EMBEDDINGS OF MANIFOLDS.*

By J. P. Hempel and D. R. McMillan, Jr.

---

1. Introduction and Definitions. In [4] and [5], O. G. Harrold showed that if N^k is a topological k-cell embedded in the n-sphere, S^n, $(k \neq n)$ in a locally "nice" manner (in the sense that $S^n - N^k$ has "uniformly abelian local fundamental groups"), then each topological cell contained in N^k has a simply-connected complement in S^n. In particular, his results apply to locally flat embeddings (see [2]), but there is the advantage that his concept of a "locally nice embedding" is more widely applicable.

The results obtained here show that some even stronger properties of subsets of manifolds can be deduced from the hypotheses studied by Harrold, using recent tools of piecewise-linear topology. The hypotheses of [4] and [5] are here weakened by permitting N^k to be any piecewise-linear (abbreviated: pwl) k-manifold and by replacing S^n with any pwl n-manifold M^n $(k \neq n$, as before). On the other hand, the stronger conclusion is drawn that for $n \geqq 5$ each compact absolute retract (abbreviated: CAR) in N^k is cellular in M^n.

The definitions of *piecewise-linear manifold, cellularity*, etc., are standard, and the reader is referred to [7] and [6] for these. As remarked in [6], if $n \neq 4$ and X is cellular in the pwl manifold M^n, then X is cellular with respect to pwl cells in M^n. Hence, only the shorter phrase "cellular" will be used in this paper. The symbol Δ^n will denote the closed n-simplex, and S^n the n-sphere. Distances in the manifold M^n will be written $d(x, y)$ and $B(X; \epsilon)$ will mean the set of points of M^n each of whose distance from X is less than ϵ.

If M^n is a piecewise-linear manifold $(n \geqq 3)$ and X is a compact subset of $\operatorname{Int} M^n$, let us say that X satisfies the *cellularity criterion* if, for each open set U such that $X \subset U \subset M^n$, there is an open set V such that $X \subset V \subset U$, the inclusion $V \rightarrow U$ is homotopic to a constant, and each loop in $V - X$ is null-homotopic in $U - X$. The second author has shown in [6; Theorem 1] that if X satisfies the cellularity criterion and $n \geqq 5$, then X is cellular in M^n. The same conclusion holds for $n = 3$ if some neighbor-

---

Received April 2, 1964.

* Research supported by grant NSF-GP779.

hood of X is embeddable in E^3. If X is a CAR, then open sets V and U can always be chosen as above so that the inclusion $V \to U$ is homotopic to a constant.

We also adapt the following notations from [3] and [4]. Let A be a subset of the separable metric space Z. If $x \in \bar{A}$, we say that A is *n-LC* at x if for each open subset U of Z containing x there is an open subset V of Z containing x such that each mapping of S^n into $V \cap A$ is null-homotopic in $U \cap A$. If A is *k-LC* at x for $0 \leqq k \leqq n$, then we say that A is *LC^n* at x. We say that A has *abelian local fundamental groups* (1-ALG) at $x \in \bar{A}$ if for each open subset T of Z containing x there are open subsets U and V of Z containing x such that $V \subset U \subset T$ and each mapping f of S^1 into $V \cap A$ which "bounds" in $U \cap A$ is null-homotopic in $U \cap A$. By saying that f "bounds," we mean that if s is the fundamental 1-cycle on S^1 and if

$$f^*: H_1(S^1) \to H_1(U \cap A)$$

is the induced homomorphism, then $f^*(s) = 0$ (singular homology with integer coefficients is understood). For example, if N^k is a locally flat k-sphere in S^n, then each component of $S^n - N^k$ is LC^1 at each point of N^k when $1 \leqq k \leqq n-1$, $k \neq n-2$; and $S^n - N^k$ has 1-ALG at each point of N^k in case $k = n-2$.

We can now state formally our main result:

THEOREM 1. *Let M^n be a piecewise-linear manifold of dimension n, $n \geqq 3$, and N^k a piecewise-linear manifold of dimension k, $k \leqq n-1$, topologically embedded in* $\operatorname{Int} M^n$. *If $k \neq n-2$, suppose that $M^n - N^k$ is* 1-LC *at each point of N^k. If $k = n-2$, suppose that $M^n - N^k$ has* 1-ALG *at each point of N^k. Let X be a compact absolute retract in N^k. Then, X satisfies the cellularity criterion. Hence, if $n \geqq 5$, X is cellular in M^n.*

The hypotheses on the embedding of N in M in Theorem 1 will be abbreviated by saying that the embedding is *locally nice.*

We note that the above theorem holds, in particular, when $X = N^k$ is a k-cell in M^n which satisfies the "locally nice" embedding assumptions (which certainly are satisfied if N^k is locally flat in M^n).

Question. It is known [6; Theorem 4] that if $S^{n-1} \subset S^n$ $(n \geqq 5)$ is an embedding such that $S^n - S^{n-1}$ is 1-LC at each point of S^{n-1} then each component of $S^n - S^{n-1}$ is an open n-cell. Is it true that if $S^k \subset S^n$ $(n \geqq 5$, $k \leqq n-3)$ is an embedding such that $S^n - S^k$ is 1-LC at each point of S^k then $S^n - S^k$ is homeomorphic to the complement of an unknotted k-sphere in S^n? An application of Thorem 1 shows the answer to be "yes" if we make the additional assumption that S^k be locally flat at some point; for choose a k-cell X in S^k such that $S^k - X$ is locally flat in S^n. Let S_*^n be the decomposition space obtained from S^n by collapsing X to a point and let $\eta\colon S^n \to S_*^n$ be the natural map. By Theorem 1 X is cellular in S^n; thus S_*^n is an n-sphere. By choosing X to be "nicely" embedded in S^k we can insure that $S_*^k = \eta(S^k)$ is a k-sphere; furthermore S_*^k will be locally flat in S_*^n except at the point $\eta(X)$. By [8; Theorem 9.2] it follows that S_*^k is unknotted in S_*^n. Since $\eta \mid S^n - S^k$ is a homeomorphism, our assertion follows.

The next two sections are devoted to the proof of Theorem 1. In Section 4, we give conditions under which products of embeddings are locally nice. In Section 5, we give an example to show that a locally nice embedding of S^{n-2} in S^n may fail to be locally flat. We do not know if there are such examples for each positive codimension other than two. It seems likely that there are.

2. Theorem 1: The case $k \leqq n-2$.

Throughout this section, we assume the hypotheses and notation of Theorem 1 as stated earlier (with $n \geqq 3$). We also suppose that $k \leqq n-2$. The case $k = n-1$ will be handled in Section 3.

First of all, if $k \leqq n-3$, a result even stronger than that stated by Theorem 1 holds. In fact, according to [6; Lemmas 2 and 3], if K is a compact set of dimension k in $\operatorname{Int} M^n$, where $n \geqq 3$ and $k \leqq n-3$, and $M^n - K$ is 1-LC at each point of K, then the cellularity criterion is satisfied for each CAR contained in K.

Now suppose that $k = n-2$. The proof in this case can then be obtained by slight modifications of the proofs found in [4; Lemma 5.1] on pages 126 and 127. We outline the modified argument here and refer the reader to [4] for details.

Let $X \subset N^{n-2}$ be the CAR for which we wish to verify the cellularity criterion, and let an open set U be given such that $X \subset U \subset M^n$. Since $\dim X \leqq n-2$, $M^n - X$ is 0-LC at each point of X, and it will suffice to show that there exists a neighborhood $V^* \subset U$ of X and $\epsilon > 0$ such that each loop of diameter less than ϵ in $V^* - X$ is null-homotopic in $U - X$. For, we may then choose an open set V such that $X \subset V \subset V^*$ and the inclusion $V \to V^*$ is homotopic to a constant. The required open set is V.

To show the existence of V^* and ϵ, we proceed essentially as in [4; Lemma 5.1]. By Lemma 1 (see the end of this section), we may assume that $X \cap \operatorname{Bd} N^{n-2} \neq \emptyset$. For each $x \in X$, there are, by the 1-ALG property, connected open sets $Q(x)$ and $W(x)$ containing x, such that $Q(x) \subset W(x) \subset U$, and each loop in $Q(x) - N^{n-2}$ which "bounds" in $W(x) - N^{n-2}$ is null-homotopic in $W(x) - N^{n-2}$. We may assume that $Q(x) \cap \mathrm{Bd}\, N^{n-2} = \emptyset$ if $x \in X \cap \mathrm{Int}\, N^{n-2}$. Note also that each loop in $Q(x) - N^{n-2}$ is null-homotopic in $W(x) - N^{n-2}$ if $x \in X \cap \mathrm{Bd}\, N^{n-2}$.

Let $Q(x_1), \cdots, Q(x_s)$ cover X. Note that $x_i \in X$ and that at least one x_i belongs to $\mathrm{Bd}\, N^{n-2}$. Choose ϵ so that each set of diameter less than ϵ in $B(X;\epsilon)$ lies in some $Q(x_i)$. Define V^* to be $B(X;\epsilon)$. Now, if a loop in $V^* - X$ of diameter less than ϵ is given, it lies in some $Q(x_i)$, and it may even be assumed to lie in $Q(x_i) - N^{n-2}$. Since X is connected, there is a finite chain of $Q(x_i)$'s, beginning with the one containing the loop, ending with a $Q(x_i)$ for which $x_i \in \mathrm{Bd}\, N^{n-2}$, and such that adjacent $Q(x_i)$'s in the chain meet in a point of $X \cap \mathrm{Int}\, N^{n-2}$. Let $Q(x_i)$ and $Q(x_j)$ be the first two elements in this chain. Then the given loop is homologous, thus by the 1-ALG property, is homotopic in $W(x_i) - N^{n-2}$ to a loop in $Q(x_i) \cap Q(x_j) - N^{n-2}$. If $Q(x_k)$ is the third element in the chain, then the new loop is in turn homotopic in $W(x_j) - N^{n-2}$ to a loop in $Q(x_j) \cap Q(x_k) - N^{n-2}$, and so on. Finally, we obtain a loop in (say) $Q(x_q) - N^{n-2}$, where $x_q \in \mathrm{Bd}\, N^{n-2}$. This last loop is then null-homotopic in $W(x_q) - N^{n-2}$, which completes the proof.

To justify the assumption that $X \cap \mathrm{Bd}\, N^{n-2} \neq \emptyset$ in the above outline, we need the following:

LEMMA 1. *Let X be a compact absolute retract in* $\mathrm{Int}\, M^n$, *where M^n is a piecewise-linear manifold of dimension n, $n \geqq 3$. Let A be an arc in* $\mathrm{Int}\, M^n$ *such that $A \cap X$ consists of a single point. Then, if $X \cup A$ satisfies the cellularity criterion, so also does X.*

Proof. We refer the reader to [6; Theorem 6], where this lemma is proved in the case that X is an arc. The proof is the same except for this change.

3. Theorem 1: The case $k = n - 1$.

In this section we assume that M is a pwl manifold of dimension $n \geqq 3$ and that N is a pwl manifold of dimension $n - 1$ which is topologically embedded in $\mathrm{Int}\, M$. We present several lemmas which lead to the proof of Proposition 1 which states that if $M - N$ is 1-LC at each point of N then each CAR in N satisfies the cellularity criterion in M.

The following is a special case of [3; Theorem 2] and is stated here for the reader's convenience.

LEMMA 2. *Let A be a subset of a separable metric space Z such that A is* LC^1 *at each point of $\bar{A} \subset Z$. Suppose we are given a map $f: \Delta^2 \to \bar{A}$, a closed subset B of Δ^2, and a positive number ϵ. Then there is a map $f^*: \Delta^2 \to \bar{A}$ such that $f^* \mid B = f \mid B$, $f^*(\Delta^2 - B) \subset A$, and $d(f(x), f^*(x)) < \epsilon$ for each $x \in \Delta^2$.*

LEMMA 3. *If $x \in \operatorname{Int} N$, then for each $\epsilon > 0$ there corresponds a $\delta > 0$ such that $B(x;\delta) - N = W_1 \cup W_2$ where W_1 and W_2 are disjoint open sets such that any pair of points of W_i $(i = 1, 2)$ can be joined by an arc in $B(x;\epsilon) - N$. Whereas if $x \in \operatorname{Bd} N$ then for each $\epsilon > 0$ there corresponds a $\delta > 0$ such that any pair of points of $B(x;\delta) - N$ can be joined by an arc in $B(x;\epsilon) - N$.*

The above lemma is a well known fact (see, for example, [1; Section 2, Corollary II]), as is the following which can be established by repeated application of the above.

LEMMA 4. *If $x \in N$, then for any $\epsilon > 0$ there corresponds a $\delta > 0$ such that any point of $B(x;\delta) - N$ can be joined to any point y of $B(x;\delta) \cap N$ by an arc which lies, except for y, in $B(x;\epsilon) - N$.*

An arc A is said to *pierce* N at a point $x \in \operatorname{Int} N$ if there is a subarc B of A such that $x \in \operatorname{Int} B$ and such that for some $\epsilon > 0$ each of the corresponding sets W_1 and W_2 of Lemma 3 contains exactly one component of $B - x$. This concept is independent of the choice of ϵ and hence is well defined.

A finite graph K (homeomorphic image of a 1-complex) is said to be in *special position* with respect to N if $K \cap N$ is a finite set in $\operatorname{Int} N$ which contains no vertex of K and such that N is pierced by some arc in K at each point of $K \cap N$.

LEMMA 5. *If K is a finite graph in M and $\epsilon > 0$ then there is a homeomorphism h of K into M which moves no point as much as ϵ and such that $h(K)$ is in special position with respect to N. Furthermore if L is any subgraph of K which is in special position with respect to N then h can be chosen to be the identity on L.*

Proof. We assume that $K \cap \operatorname{Bd} N = \emptyset$; for this can be accomplished by a slight adjustment of K which leaves L fixed. Similarly we may assume that no vertex of K lies in N. For each $x \in K \cap N$ let

$$\eta(x) = \min(\epsilon/3, d(x, \operatorname{Bd} N)).$$

Corresponding to each $\eta(x)$ take $\delta(x) > 0$ small enough to satisfy the requirements of both Lemma 3 and Lemma 4. There is a finite set $x_1, \cdots, x_n$ of points of $K \cap N$ such that the collection $\{B(x_i; \delta(x_i))\}$ covers $K \cap N$. As assured by Lemma 3 we have $B(x_i; \delta(x_i)) - N = W_1(x_i) \cup W_2(x_i)$.

There is a homeomorphism h_1 of K into M which is the identity on L, which moves no point as much as $\epsilon/3$, and such that $h_1(K - L) \cap N$ is contained in the union of the interiors of a finite collection $\{A_j\}$ of mutually exclusive arcs in $h_1(K - L)$ each of diameter less than $\epsilon/3$ and each lying in some set $B(x_i; \delta(x_i))$. If the endpoints of A_j lie in the same $W_k(x_i)$ we can, by Lemma 3, join them by an arc B_j which misses N and has diameter less than $\epsilon/3$. If, on the other hand, one endpoint of A_j lies in $W_1(x_i)$ while the other endpoint lies in $W_2(x_i)$, we can, by Lemma 4, join these points by an arc B_j of diameter less than $\epsilon/3$ which intersects N at a single point at which the arc pierces N. We choose the B_j's so that their interiors are disjoint and so that none of them intersects $h_1(K) - \cup A_j$. We define $h_2\colon h_1(K) \to M$ to be the identity on $h_1(K) - \cup A_j$ and to take A_j onto B_j. The desired homeomorphism is $h = h_2 h_1$.

LEMMA 6. *Suppose that J is a simple closed curve in $M - \mathrm{Bd}\,N$ which is in special position with respect to N and suppose that J is contractible in $M - \mathrm{Bd}\,N$. Then $J \cap N$ must consist of an even number of points.*

Proof. Suppose J satisfies the hypothesis of our lemma. Then there is a map f of Δ^2 into $M - \mathrm{Bd}\,N$ such that $f \mid \mathrm{Bd}\,\Delta^2$ is a homeomorphism of $\mathrm{Bd}\,\Delta^2$ onto J. If for each $x \in f(\Delta^2) \cap N$ we let $\epsilon(x) = d(x, \mathrm{Bd}\,N)$, there corresponds by Lemma 3 a number $\delta(x) > 0$ such that $B(x; \delta(x)) - N = W_1(x) \cup W_2(x)$ where the sets $W_i(x)$ satisfy the conclusions of Lemma 3. Now let T be a triangulation of Δ^2 of small enough mesh that if $f(\sigma_j) \cap N \neq \emptyset$ then $f(\sigma_j)$ lies in one of the sets $B(x; \delta(x))$, for some $x \in f(\Delta^2) \cap N$ (where $\{\sigma_j\}$ denotes the collection of 2-simplexes of T). We can approximate $f \mid T^1$ (where T^1 is the 1-skeleton of T) by a homeomorphism and then by Lemma 5 further approximate to obtain a homeomorphism g of T^1 into $M - \mathrm{Bd}\,N$ such that $g \mid \mathrm{Bd}\,\Delta^2 = f \mid \mathrm{Bd}\,\Delta^2$ and such that $g(T^1)$ is in special position with respect to N. We choose g close enough to f so that if $g(\mathrm{Bd}\,\sigma_j) \cap N \neq \emptyset$ then $g(\mathrm{Bd}\,\sigma_j)$ must lie in $B(x; \delta(x))$ for some $x \in f(\Delta^2) \cap N$. Now $\#(g(\mathrm{Bd}\,\sigma_j), N) \equiv 0 \bmod 2$ (where $\#(A, B)$ denotes the number of points in $A \cap B$); for if the image of some edge of σ_j intersects N in an odd number of points this must be true of exactly two edges of σ_j. Thus

$$\sum_j \#(g(\mathrm{Bd}\,\sigma_j), N) \equiv 0 \bmod 2.$$

Since each "interior" 1-simplex of T^1 occurs twice in this sum, we have

$$0 \equiv \sum_j \#(g(\mathrm{Bd}\,\sigma_j), N) \equiv \sum_j (g(\mathrm{Bd}\,\sigma_j \cap \mathrm{Bd}\,\Delta^2), N)$$
$$\equiv \#(g(\mathrm{Bd}\,\Delta^2), N) \equiv \#(J, N) \bmod 2,$$

and the proof is complete.

LEMMA 7. *Suppose* $M - N$ *is* 1-LC *at each point of* N *and* $x \in \mathrm{Int}\,N$. *Then for each positive number* ϵ *there corresponds a positive number* δ *such that if* f *is a map of* $\mathrm{Bd}\,\Delta^2$ *into* $B(x;\delta)$ *with the property that* $f^{-1}(N)$ *is an arc or a point, then* f *can be extended to a map* F *of* Δ^2 *into* $B(x;\epsilon)$ *such that* $F(\mathrm{Int}\,\Delta^2) \cap N = \emptyset$.

Proof. Given $x \in \mathrm{Int}\,N$ and $\epsilon > 0$, there is by Lemma 3 a positive number η such that $B(x;\eta) - N$ is the union of two nonempty open sets W_1 and W_2 which satisfy the conclusion of Lemma 3. There is a positive number θ such that $\rho(B(x;\theta)) \subset B(x;\eta) \cap N$, where ρ is a retraction of some neighborhood of N onto N. Finally choose $\delta > 0$ so that $B(x;\delta)$ is contractible in $B(x;\theta)$.

Now suppose f is a map of $\mathrm{Bd}\,\Delta^2$ into $B(x;\delta)$ such that $f^{-1}(N)$ is an arc or a point. Then f can be extended to a map g of Δ^2 into $B(x;\theta)$. Now there is a value of i such that $f(\mathrm{Bd}\,\Delta^2) \subset N \cup W_i$ (assume for convenience that $i = 1$). Define $G\colon \Delta^2 \to W_1 \cup (N \cap B(x;\eta))$ by $G(x) = g(x)$ if $x \in g^{-1}(W_1)$ and $G(x) = \rho g(x)$ otherwise. Note that $G \mid \mathrm{Bd}\,\Delta^2 = f$. Now it follows from Lemma 3 together with the assumption that $M - N$ is 1-LC at each point of N that W_1 is LC^1 at each point of $W_1 \cup (N \cap B(x;\eta))$. Thus there is, by Lemma 2, a map F of Δ^2 into $W_1 \cup (N \cap B(x;\eta))$ such that $F \mid \mathrm{Bd}\,\Delta^2 = G \mid \mathrm{Bd}\,\Delta^2 = f$ such that $F(\mathrm{Int}\,\Delta^2) \subset W_1$. This completes the proof of our lemma.

LEMMA 8. *Suppose* $M - N$ *is* 1-LC *at each point of* N *and* $x \in \mathrm{Bd}\,N$. *Then for each positive number* ϵ *there corrresponds a positive number* δ *such that if* f *is a map of* $\mathrm{Bd}\,\Delta^2$ *into* $B(x;\delta)$ *with the property that* $f^{-1}(N)$ *is an arc and* $f^{-1}(\mathrm{Bd}\,N)$ *is a single point, then* f *can be extended to a map* F *of* Δ^2 *into* $B(x;\epsilon)$ *such that* $F(\mathrm{Int}\,\Delta^2) \cap N = \emptyset$.

Proof. Given $x \in \mathrm{Bd}\,N$ and $\epsilon > 0$, there is a number $\delta > 0$ such that any loop in $B(x;\delta) - N$ is null homotopic in $B(x;\epsilon) - N$. Now suppose f is a map of $\mathrm{Bd}\,\Delta^2$ into $B(x;\delta)$ such that $f^{-1}(N)$ is an arc and $f^{-1}(\mathrm{Bd}\,N)$ is a single point. We assume that $f^{-1}(N)$ is the base, B, of Δ^2 and $f^{-1}(\mathrm{Bd}\,N)$ is the midpoint, p, of this base (for the case in which $f^{-1}(\mathrm{Bd}\,N)$ is an endpoint of $f^{-1}(N)$ the proof is essentially the same). Let

$$\epsilon_1 = \min[\epsilon/2, d(f(p), M - B(x;\epsilon))].$$

There is a number $\delta_1 > 0$ such that any loop in $B(f(p);\delta_1) - N$ is null homotopic in $B(f(p);\epsilon_1) - N$. For any $y \in f(B) - f(p)$ there is, by Lemma 7, a number $\delta(y) > 0$ such that if g is a map of the boundary of a disk into $B(y;\delta(y))$ such that $g^{-1}(N)$ is an arc or a point then g can be extended to a map of the disk into $B(x;\epsilon)$ which takes no interior point of the disk into N.

We choose a cell decomposition T of Δ^2 as follows. Let Δ_0 be obtained from Δ^2 by raising its base slightly and let the "strip" $\Delta^2 - \Delta_0$ be triangulated into several 2-cells Δ_i $(i \geqq 1)$ with p the midpoint of the base of Δ_1 (refer to Figure 1(a)). By use of Lemmas 3 and 4 T can be constructed so that there is an extension f_1 of f to the 1-skeleton of T with the properties that $f_1^{-1}(N) = B$, $f_1(\operatorname{Bd}\Delta_1) \subset B(f(p);\delta_1)$, and for each $i \geqq 2$ there exists $y \in f(B) - f(p)$ such that $f_1(\operatorname{Bd}\Delta_i) \subset B(y;\delta(y))$. By choice of δ, $f_1 \mid \operatorname{Bd}\Delta_0$ can be extended to map Δ_0 into $B(x;\epsilon) - N$ and by the choice of the $\delta(y)$'s, $f_1 \mid \operatorname{Bd}\Delta_i$ $(i \geqq 2)$ can be extended to map Δ_i into $B(x;\epsilon)$ in such a way that no interior point of Δ_i is mapped into N. The net effect is to extend f to a map F_1 of $\Delta^2 - \operatorname{Int}\Delta_1$ into $B(x;\epsilon)$ such that $F_1^{-1}(N) = B$. By treating Δ_1 as we just treated Δ^2 we can extend F_1 to a map F_2 of $\Delta^2 - \operatorname{Int}\Delta_{11}$ into $B(x;\epsilon)$ where Δ_{11} is a small 2-cell in Δ_1 which contains p in its base (see Figure 1(b)). The desired map F will be the limit of a sequence $\{F_i\}$ of maps described as above.

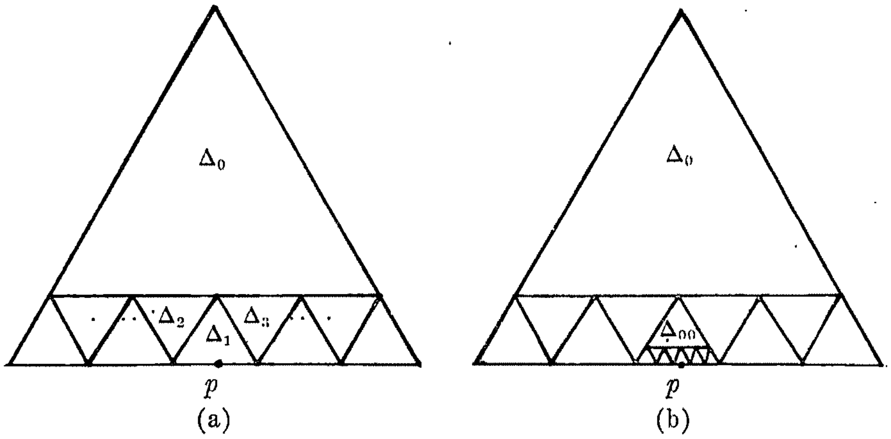


FIGURE 1.

LEMMA 9. *Suppose that $M - N$ is 1-LC at each point of N and that X is a compact absolute retract contained in N. Let U be an open set in M containing X. Then there is a compact piecewise-linear $(n-1)$-manifold N_1*

with connected, non-empty boundary such that $X \subset N_1 \subset N \cap U$ *and such that* $M - N_1$ *is* 1-LC *at each point of* N_1.

Proof. First, we outline a proof of the following fact: If X is a CAR contained in the pwl manifold N^k and V is an open subset of N^k containing X, then there is a compact pwl manifold N_1^k such that N_1^k is a polyhedron in N^k, $X \subset N_1^k \subset V$, and $\operatorname{Bd} N_1^k$ is connected (note: possibly $X \cap \operatorname{Bd} N^k \neq \emptyset$).

To show this, we first choose compact pwl k-manifolds N_2^k and N_3^k (with non-empty boundary) such that $X \subset N_3^k \subset N_2^k$, and each of the inclusions $N_3^k \to N_2^k$, $N_2^k \to V$ is homotopic to a constant. Let F^{k-1} be a component of $\operatorname{Bd} N_3^k$. Since the fundamental homology cycle carried by F^{k-1} is a boundary in N_2^k, it follows that there is a unique component K of $N_2^k - F^{k-1}$ (whose closure we call the *fundamental region* of F^{k-1}) having the property that $\bar{K}$ is a compact pwl manifold with $\operatorname{Bd} \bar{K} = F^{k-1}$. Note that $N_2^k - F^{k-1}$ is connected if and only if $F^{k-1} \subset \operatorname{Bd} N_2^k$. If the fundamental region of F^{k-1} contains X, we say that F^{k-1} *encloses* X. Now, to obtain the desired N_1^k, we begin with N_3^k and adjoin to it the fundamental region of each component of $\operatorname{Bd} N_3^k$ that fails to enclose X. This yields a compact pwl manifold N_0^k with non-empty boundary and $X \subset N_0^k \subset N_2^k$. Further, each component of $\operatorname{Bd} N_0^k$ encloses X. The fundamental region of a component of $\operatorname{Bd} N_0^k$ (actually, there is only one, since N_0^k is connected) is taken to be N_1^k. This completes the proof of the first assertion.

The above allows us to choose for the N_1 of Lemma 9 a compact pwl $(n-1)$-manifold N_1 with connected, non-empty boundary such that $X \subset N_1 \subset N \cap U$, and N_1 is polyhedral relative to some pwl structure on N. It remains to show that $M - N_1$ is 1-LC at each point of N_1. Clearly, $M - N_1$ is 1-LC at each point of $\operatorname{Int} N_1$.

Suppose first that $x \in (\operatorname{Bd} N_1) \cap (\operatorname{Int} N)$ and let $\epsilon > 0$ be given. By Lemma 7 choose $\eta > 0$ so that any map of $\operatorname{Bd} \Delta^2$ into $B(x;\eta)$ whose inverse image of N is an arc can be extended to a map of Δ^2 into $B(x;\epsilon)$ which takes $\operatorname{Int} \Delta^2$ into $B(x;\epsilon) - N$. Take $\delta > 0$ so that any pair of points of $B(x;\delta) \cap (N - N_1)$ can be joined by an arc in $B(x;\eta) \cap (N - N_1)$. We assume that $\delta < \eta < \epsilon$. Now let s_1 be a loop in $B(x;\delta) \cap (M - N_1)$. By Lemma 6 s_1 is homotopic in $B(x;\delta) \cap (M - N_1)$ to a nonsingular loop s_2 which is in special position relative to N. By joining consecutive points of $s_2 \cap N$ by arcs in $B(x;\eta) \cap (N - N_1)$, we see that s_2 is homotopic in $B(x;\eta) \cap (M - N_1)$ to a "product" $r_1 r_2 \cdots r_k$ of loops each based at the same point of $N - N_1$, and each of which meets N in an arc. Now by Lemma 7 each r_i is null-homotopic in $B(x;\epsilon) - N_1$.

The proof in case $x \in (\operatorname{Bd} N_1) \cap (\operatorname{Bd} N)$ is similar except that Lemma 8 is used in place of Lemma 7.

LEMMA 10. *Suppose X is a compact absolute retract in N and U is an open set containing X. Then there is an $(n-1)$-manifold N_1 with connected boundary such that $X \subset N_1 \subset N \cap U$ and there is an open set Q such that* $\operatorname{Int} N_1 \subset Q \subset U$ *and such that $Q - N_1$ is the union of two disjoint, nonempty open sets Q_1 and Q_2 such that $\bar{Q}_1 \cap \bar{Q}_2 = \bar{Q}_i \cap N_1 = N_1$ and such that Q_i is* 0-LC *at each point of* $\operatorname{Int} N_1$. *If, in addition, we assume that $M - N$ is* 1-LC *at each point of N then N_1 can be chosen so that $M - N_1$ is* 1-LC *at each point of N_1 and each of the sets Q_i will be* LC^1 *at each point of* $\operatorname{Int} N_1$.

Proof. Since X is an absolute retract (and hence contractible), we can find an $(n-1)$-manifold N_1 in N which contains X and lies in a small enough neighborhood of X so that N_1 is contractible in N. As in the proof of Lemma 9 we can take N_1 to have connected boundary. If we assume that $M - N$ is 1-LC at each point of N, then Lemma 9 tells us that we can choose N_1 so that $M - N_1$ is 1-LC at each point of N_1.

Since N_1 is a compact absolute neighborhood retract, there is an open set G containing N_1 and a retraction π of G onto N_1. We take a sequence $\{C_i\}$ of compact connected subsets of $\operatorname{Int} N_1$ such that $C_i \subset C_{i+1}$ and $\operatorname{Int} N_1 = \bigcup_i C_i$. For each i let H_i be an open set in M such that $C_i \subset H_i \subset G$ and $\pi \mid H_i$ is homotopic to the identity in $M - (N - \operatorname{Int} N_1)$.

Now for each $x \in C_i$ and fixed i let $\epsilon(x) = d(x, M - H_i)$. Corresponding to each $\epsilon(x)$ there is a number $\delta(x) > 0$ which satisfies the requirements of both Lemma 3 and Lemma 4. Then $B(x; \delta(x)) - N_1 = W_1(x) \cup W_2(x)$ where the sets $W_j(x)$ satisfy the conclusion of Lemma 3. We let

$$Q^i = \bigcup_{x \in C_i} B(x; \delta(x)).$$

Pick $x_0 \in C_i$ and let $Q_1{}^i$ be the union of all those sets $W_j(x)$, $x \in C_i$, for which there is an arc in $H_i - N$ joining a point of $W_j(x)$ to a point of $W_1(x_0)$. Now for each $x \in C_i$ at least one of the sets $W_1(x)$, $W_2(x)$ must be contained in $Q_1{}^i$; for the set of all points of C_i for which this is true is both open and closed in C_i. We assume the notation is chosen so that for each $x \in C_i$, $W_1(x) \subset Q_1{}^i$. Furthermore, for each $x \in C_i$, $W_2(x) \cap Q_1{}^i = \emptyset$; for suppose not. Then there is a point $x_1 \in C_i$ such that $W_2(x_1) \cap Q_1{}^i \neq \emptyset$. Thus there is a point $x_2 \in C_i$ and there is a value of j such that $W_2(x_1) \cap W_j(x_2) \neq \emptyset$ and such that there is an arc in $H_i - N$ joining a point of $W_1(x_0)$ to a point of $W_j(x_2)$. By an application of Lemma 3, this arc can be extended to an arc A in $H_i - N$ which joins a point a_1 of $W_1(x_0)$ to a point a_2 of $W_2(x_1) \cap W_j(x_2)$. By a similar argument, together with the fact that $W_1(x_1) \subset Q_1{}^i$, there is an arc B in $H_i - N$ joining a point b_1 of $W_1(x_0)$ to a point b_2 of $W_1(x_1)$. By Lemma 3 there is an arc C from a_1 to b_1 in $H_i - N$, and by two applications of Lemma 4 there is an arc D from a_2 to b_2 which pierces N at a point of $B(x_1; \delta(x_1)) \cap N$ and otherwise misses N. We choose A, B, C, and D so that their interiors are disjoint. Thus the simple closed curve $J = A \cup B \cup C \cup D$ in H_i is in special position with respect to N and intersects N in a single point. However $\pi \mid J$ is homotopic to the identity in $M - \operatorname{Bd} N$ and, since $\pi(J) \subset \operatorname{Int} N_1$, $\pi \mid J$ is homotopic in N to a constant map. Since $\operatorname{Bd} N$ has a cartesian product neighborhood in N and $\pi(J) \subset \operatorname{Int} N$, $\pi \mid J$ is homotopic to a constant map in $\operatorname{Int} N$. Thus we have shown that J is contractible in $M - \operatorname{Bd} N$. This contradicts Lemma 6 and, therefore, establishes our assertion that no $W_2(x)$ meets $Q_1{}^i$. We let $Q_2{}^i = \bigcup_{x \in C_i} W_2(x)$. Then $Q^i - N$ is the union of the nonempty disjoint sets $Q_1{}^i$ and $Q_2{}^i$. By another application of the "linking argument" (Lemma 6) it can be shown that for any i and j each of the sets $Q_1{}^i$ and $Q_2{}^i$ has a nonempty intersection with exactly one of the sets $Q_1{}^j$ and $Q_2{}^j$. We assume the notation chosen so that $Q_1{}^i \cap Q_1{}^j \neq \emptyset$. We now let $Q = \bigcup_i Q^i$ and $Q_k = \bigcup_i Q_k{}^i$ $(k = 1, 2)$. Then $\operatorname{Int} N_1 \subset Q \subset U$ and $Q - N_1$ is the union of the nonempty, disjoint open sets Q_1 and Q_2. It can be easily verified that the sets Q_1 and Q_2 satisfy the remaining requirements of our lemma.

PROPOSITION 1. *Suppose N is a piecewise-linear $(n-1)$-manifold, $n \geqq 3$, which is topologically embedded in the interior of a piecewise-linear n-manifold M in such a way that $M - N$ is* 1-LC *at each point of N. If X is a compact absolute retract in N then for any open set U containing X there is an open set V containing X such that any simple closed curve in $V - X$ is contractible in $U - X$.*

Proof. Let U be an open set containing X. We choose an open set $V (X \subset V \subset U)$ so that V is contractible in U. We wish to show that V satisfies the conclusion of our theorem. To this end let J be a simple closed curve in $V - X$. Since J is contractible in U, there is a map $f: \Delta^2 \to U$ such that $f \mid \operatorname{Bd} \Delta^2$ is a homeomorphism of $\operatorname{Bd} \Delta^2$ onto J. We will show that f can be replaced by a map $f^*: \Delta^2 \to U - X$ such that $f \mid \operatorname{Bd} \Delta^2 = f^* \mid \operatorname{Bd} \Delta^2$.

By Lemma 10 there is an $(n-1)$-manifold N_1 with connected boundary such that $X \subset N_1 \subset N \cap (V - J)$ and there is an open set Q such that

$\operatorname{Int} N_1 \subset Q \subset V$ and such that $Q - N_1$ is the union of two nonempty, disjoint open sets Q_1 and Q_2 such that $\bar{Q}_1 \cap \bar{Q}_2 = \bar{Q}_i \cap N_1 = N_1$ and such that Q_i is LC^1 at each point of $\operatorname{Int} N_1$. By Lemma 9 there is an $(n-1)$-manifold N_2 with connected boundary such that $X \subset N_2 \subset N_1$, such that N_2 is contractible in N_1, and such that $M - N_2$ is 1-LC at each point of N_2.

Our first step will be to replace f by a map $g\colon \Delta^2 \to U$ such that $g \mid \operatorname{Bd} \Delta^2 = f \mid \operatorname{Bd} \Delta^2$ and $g(\Delta^2) \cap \operatorname{Int} N_2 = \emptyset$.

We can, by application of Lemmas 7 and 8, find a finite collection $\{R_1, \cdots, R_k\}$ of open sets which covers $f(\Delta^2) \cap N_2$ and such that if h is a map of the boundary of a disk D into R_i such that $h^{-1}(N_2)$ is an arc and, in case $R_i \cap \operatorname{Bd} N_2 \neq \emptyset$ $h^{-1}(\operatorname{Bd} N_2)$ is exactly one point, then h can be extended to a map H of D into U such that $H(\operatorname{Int} D) \cap N_2 = \emptyset$. There is no loss of generality in assuming that $R_i \cap N_2$ is connected. There is a collection $\{S_\alpha\}$ of open sets which covers $f(\Delta^2) \cap N_2$ and which is a star refinement of $\{R_i\}$ (i. e. for each α, the union of all elements of the collection which intersect S_α lies in some R_i). We now take, by use of Lemma 4, a collection $\{T_\beta\}$ of open sets which cover $f(\Delta^2) \cap N_2$ such that for each β there is an α so that if $x \in T_\beta - N_2$ and $y \in T_\beta \cap N_2$ then there is an arc in S_α from x to y which intersects N_2 only at y.

Now since $N_2 \cap J = \emptyset$, there is a finite collection $\{D_1, \cdots, D_p\}$ of mutually exclusive disks in $\operatorname{Int} \Delta^2$ such that $f(\Delta^2 - \bigcup \operatorname{Int} D_i) \cap N_2 = \emptyset$ and such that for each i, $f(\operatorname{Bd} D_i) \subset \bigcup_\beta T_\beta$. Let us suppose, for simplicity, that there is only one D_i which we will denote by D—the general case is no more difficult since the procedure will be to modify the map f on each D_i. There is an annulus A in D one of whose boundary components is $\operatorname{Bd} D$. We suppose that A is identified with $(\operatorname{Bd} D) \times [0,1]$ with $\operatorname{Bd} D \times \{0\}$ corresponding to $\operatorname{Bd} D$. Denote the closure of $D - A$ by D'. We give $\operatorname{Bd} D$ a triangulation of mesh small enough that the image under f of each 1-cell of the triangulation lies in some T_β. We wish to extend $f \mid \operatorname{Bd} D$ to a map h of A into U so that $h^{-1}(N_2) = \operatorname{Bd} D'$ and $h^{-1}(\operatorname{Bd} N_2)$ is finite. We proceed as follows (refer to Figure 2). For each vertex v of our triangulation of $\operatorname{Bd} D$ we define h to map the arc $v \times [0,1]$ onto an arc from $f(v)$ to a point of $\operatorname{Int} N_2$. By our choice of the collection $\{T_\beta\}$, h may be defined on $v \times [0,1]$ so that $h(v \times [0,1))) \cap N_2 = \emptyset$ and so that $h(v \times [0,1])$ lies in some S_α. Now suppose τ is a 1-cell in our triangulation of $\operatorname{Bd} D$. Since $h(\tau) = f(\tau)$ lies in some T_β and hence in some S_α, it follows by the star refinement requirement on $\{S_\alpha\}$ that $h(\tau \cup (\operatorname{Bd} \tau) \times [0,1])$ lies in some R_i. We extend h to map $\tau \times 1$ onto an arc in $R_i \cap N_2$. In the case that $R_i \cap \operatorname{Bd} N_2 \neq \emptyset$ we require that this arc intersect Bd N_2 in exactly one point. Now we have defined h to map Bd$(\tau \times [0,1])$ into R_i in such a way that $h^{-1}(N_2) = \tau \times 1$. It therefore follows by our choice of the collection $\{R_i\}$ that h can be extended to map $\tau \times [0,1]$ into U in such a way that $h(\operatorname{Int}(\tau \times [0,1])) \cap N_2 = \emptyset$. By repeating this process for each 1-cell in Bd D we define the desired map h on all of A.

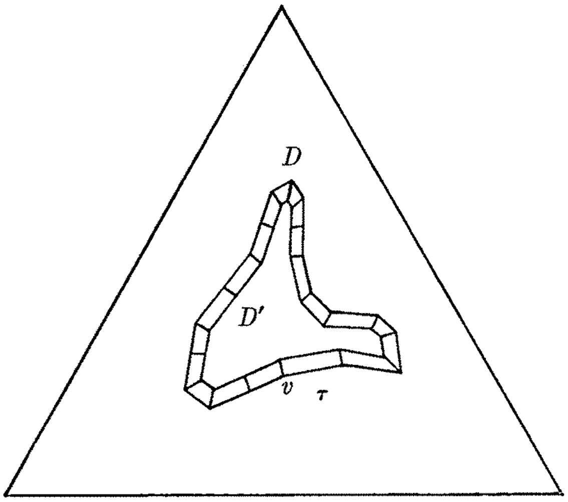


FIGURE 2.

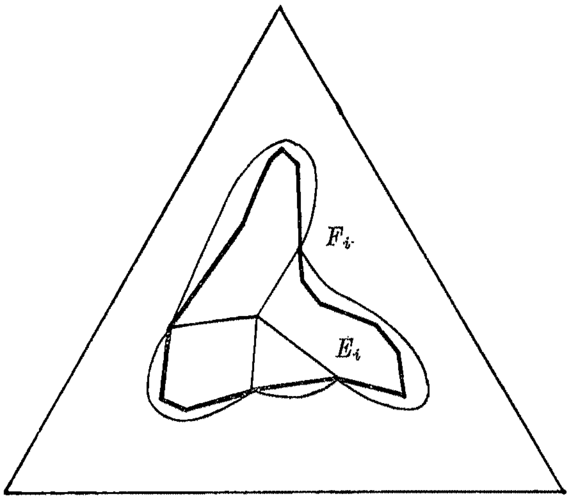


FIGURE 3.

We now join each point of $h^{-1}(\mathrm{Bd}\,N_2)$ by arc to a fixed point of $\mathrm{Ind}\,D'$ (see Figure 3). Since $\mathrm{Bd}\,N_2$ is connected, we extend h to map the union of these arcs into $\mathrm{Bd}\,N_2$. These arcs are chosen, as shown in Figure 3, to divide D' into a finite collection $\{E_i\}$ of "pie shaped" disks (we are assuming that $h^{-1}(\mathrm{Bd}\,N_2)$ contains at least two points; the following argument can be easily modified if this is not the case). Now N_2 is contractible in N_1 and for each i $h(\mathrm{Bd}\,E_i) \subset N_2$; thus we can extend h to map each E_i into N_1. We continue to denote the extension by h.

Recall that we previously chose an open set Q containing $\mathrm{Int}\,N_1$ such that $Q - N_1$ is the disjoint union of two open sets Q_1 and Q_2. Since the interior of each arc $\mathrm{Bd}\,E_i \cap \mathrm{Bd}\,D'$ is mapped by h into $\mathrm{Int}\,N_1$, we can, by "swelling" E_i slightly into A, obtain a disk F_i such that $F_i \cap D' = E_i$ and $h(F_i - E_i)$ lies entirely in one of the sets Q_1 or Q_2, say $Q_{k(i)}$. Now let $B_i = \mathrm{Bd}\,F_i \cup (h^{-1}(\mathrm{Bd}\,N_1) \cap F_i)$. Note that $h \mid F_i$ maps F_i into $Q_{k(i)} \cup N_1$. We would like to use Lemma 2 to replace $h \mid F_i$ by a map $g_i : F_i \to U$ such that $g_i \mid B_i = h \mid B_i$ and such that $g_i(F_i - B_i) \cap N_1 = \emptyset$. We are prevented from doing this directly since we do not know that $Q_{k(i)}$ is 1-LC at points of $\mathrm{Bd}\,N_1$. We proceed as follows. For each i take two sequences $\{K_j{}^i\}$ and $\{L_j{}^i\}$, $j = 1, 2, \cdots$, of (not necessarily connected) 2-manifolds with boundary such that

$$K_j{}^i \subset \mathrm{Int}\,L_j{}^i \subset L_j{}^i \subset \mathrm{Int}\,K_{j+1}{}^i$$

and such that $F_i - B_i = \bigcup_j L_j{}^i$. Now h maps $L_1{}^i$ into $Q_{k(i)} \cup \mathrm{Int}\,N_1$. Since $Q_{k(i)}$ is LC^1 at each point of $\mathrm{Int}\,N_1$ we can, by Lemma 2, replace $h \mid F_i$ by a map $g_{i1} : L_1{}^i \to Q_{k(i)} \cup \mathrm{Int}\,N_1$ such that $g_{i1} \mid \mathrm{Bd}\,L_1{}^i = h \mid \mathrm{Bd}\,L_1{}^i$ and such that $g_{i1}(\mathrm{Int}\,L_1{}^i) \subset Q_{k(i)}$. We suppose that g_{i1} is defined on all of F_i by putting $g_{i1} \mid F_i - L_1{}^i = h \mid F_i - L_1{}^i$. Now let $\epsilon_2 = \min\{1/2, \frac{1}{2}d(g_{i1}(K_1{}^i), N_1)\}$. In the same manner as above we obtain a map $g_{i2} : F_i \to Q_{k(i)} \cup N_1$ such that $g_{i2} \mid F_i - L_2{}^i = h \mid F_i - L_2{}^i$, $g_{i2}(\mathrm{Int}\,L_2{}^i) \subset Q_{k(i)}$, and g_{i2} is an ϵ_2-approximation to g_{i1}. We let $\epsilon_3 = \min\{1/4, \frac{1}{4}d(g_{i1}(K_1{}^i), N_1), \frac{1}{2}d(g_{i2}(K_2{}^i), N_1)\}$ and obtain a map $g_{i3} : F_i \to Q_{k(i)} \cup N_1$ such that $g_{i3} \mid F_i - L_3{}^i = h \mid F_i - L_3{}^i$, $g_{i3}(\mathrm{Int}\,L_3{}^i) \subset Q_{ki}$, and g_{i3} is an ϵ_3-approximation to g_{i2}. We proceed inductively to define a sequence $\{g_{ij}\}$ of maps as above and put $g_i = \lim_{j \to \infty} g_{ij}$. It is clear that g_i is continuous and that $g_i \mid B_i = h \mid B_i$. Now suppose $x \in F_i - B_i$. Then there is a j such that $x \in K_j{}^i$. Then $g_{ij}(x) \in Q_{k(i)}$; furthermore our sequence $\{\epsilon_j\}$ was chosen so that

$$d(g_i(x), g_{ij}(x)) < \sum_{j \leqq p} \epsilon_p \leqq \sum_{q=1}^{\infty} \frac{1}{2^q} d(g_{ij}(K_j{}^i), N_1) \leqq d(g_{ij}(x), N_1).$$

Hence $g_i(F_i - B_i) \subset Q_{k(i)}$. We now define the map $g: \Delta^2 \to U$ by

$$g(x) = \begin{cases} f(x)\,; x \in \Delta^2 - D \\ h(x)\,; x \in D - \bigcup_i F_i \\ g_i(x)\,; x \in F_i. \end{cases}$$

Note that $g \mid \operatorname{Bd} \Delta^2 = f \mid \operatorname{Bd} \Delta^2$ and $g(\Delta^2) \cap \operatorname{Int} N_1 = \emptyset$; so, in particular, $g(\Delta^2) \cap \operatorname{Int} N_2 = \emptyset$ as desired. The remainder of our proof is a direct application of Lemma 2. Let $U - \operatorname{Int} N_2$ correspond to the set Z of Lemma 2 and let $U - N_2$ correspond to the set A. It follows that $U - \operatorname{Int} N_2$ is LC^1 at each point of $U - N_2$; therefore there is a map $f^*: \Delta^2 \to U - \operatorname{Int} N_2$ such that $f^* \mid \operatorname{Bd} \Delta^2 = g \mid \operatorname{Bd} \Delta^2 = f \mid \operatorname{Bd} \Delta^2$ and $f^*(\operatorname{Int} \Delta^2) \subset U - N_2$. Since $f(\operatorname{Bd} \Delta^2) \subset U - N_2$, $f^*(\Delta^2) \subset U - N_2 \subset U - X$, and our proof is complete.

4. Product embeddings of manifolds.

We show in this section that, under reasonably mild restrictions, the product of two topologically embedded pwl manifolds gives a pwl manifold which is embedded in the product of the ambient manifolds in a locally nice manner.

THEOREM 2. *Let N^k and P^r be compact piecewise-linear manifolds topologically embedded in $\operatorname{Int} M^n$ and $\operatorname{Int} Q^s$, respectively, where M^n and Q^s are piecewise-linear manifolds and $n \geqq 1$, $s \geqq 1$, $n + s \geqq 3$. Then the embedding of $N \times P$ in $M \times Q$ is locally nice in each of the following situations:*

(a) $r \leqq s - 1$ *and* $k \leqq n - 2$;

(b) $r = s$, $k \leqq n - 1$, *and the embedding of N in M is locally nice*;

(c) $r = s - 1$, $k = n - 1$ *and at least one of the embeddings is locally nice.*

Proof. Case (a) is a corollary of [6; Theorem 7]. In fact, one need not even require that N and P be manifolds in this case.

Consider case (b). It suffices to show that $M \times Q - N \times P$ has 1-ALG at each point of $N \times P$, assuming that $M - N$ has 1-ALG at each point of N, since by local homology considerations, this will imply that $M \times Q - N \times P$ is 1-LC at each point of $N \times P$ in case $k \neq n - 2$. Suppose first that $(x, y) \in N \times \operatorname{Int} P$ and that the open set $A \times B$ in $M \times Q$ containing (x, y) is given. Let D^s be an open s-cell such that $y \in D^s \subset B \cap P$ and let U and V be open subsets of M such that $x \in V \subset U \subset A$ and each loop in $V \cap (M - N)$ which bounds in $U \cap (M - N)$ is null-homotopic in $U \cap (M - N)$.

Clearly,

$$(x, y) \in V \times D^s \subset U \times D^s \subset A \times B,$$

and we claim that each loop c in $(V \times D^s) \cap (M \times Q - N \times P)$ which bounds in $(U \times D^s) \cap (M \times Q - N \times P)$ is null-homotopic there. Let π_M and π_Q be the projections of $M \times Q$ onto its coordinate factors. Since $D^s \subset P$, $\pi_M(c)$ is a loop in $V \cap (M - N)$ which bounds in $U \cap (M - N)$. Hence, c is homotopic in $(U \times D^s) \cap (M \times Q - N \times P)$ to a loop c' such that $\pi_M(c') = p \in U \cap (M - N)$. Since $\pi_Q(c') \subset D^s$, c' is homotopic to a constant in $(U \times D^s) \cap (M \times Q - N \times P)$ keeping its M-coordinate fixed. This completes the proof of the special case of (b).

Now let $(x, y) \in N \times \operatorname{Bd} P$ and $k = n - 1$ in case (b). We must show that $M \times Q - N \times P$ is 1-LC at (x, y). Let K be a compact pwl s-manifold in Q such that $P \subset \operatorname{Int} K$. Then $M \times Q - N \times K$ is 1-LC at (x, y) and in fact at each point of $N \times \operatorname{Int} K$ by the special case of (b), and $N \times P \subset N \times K$. Hence we can appeal to the latter part of the proof of Lemma 9 to complete the proof of (b) when $k = n - 1$.

Now let $(x, y) \in N \times \operatorname{Bd} P$ and $k = n - 2$ in case (b). We must show that $M \times Q - N \times P$ is 1-ALG at (x, y) or equivalently, since $(x, y) \in \operatorname{Bd}(N \times P)$, that $M \times Q - N \times P$ is 1-LC at (x, y). As before, let K be a compact pwl s-manifold in Q such that $P \subset \operatorname{Int} K$. Then $M \times Q - N \times K$ is 1-ALG at (x, y) and in fact at each point of $N \times \operatorname{Int} K$ by the special case of (b). Let U be a given open set in $M \times Q$ containing (x, y) and let V and W be open sets in $M \times Q$ such that $(x, y) \in W \subset V \subset U$ and each loop in $W \cap (M \times Q - N \times K)$ which bounds in $V \cap (M \times Q - N \times K)$ is null-homotopic there. Now let c be a loop in $W \cap (M \times Q - N \times P)$. Since $\dim(N \times K) \leqq n - 2$, it may be assumed that c misses $N \times K$. But c is homologous, and hence homotopic, in $V \cap (M \times Q - N \times K)$ to a loop c' which is so close to a point of $W \cap (N \times K - N \times P)$ that c' is null-homotopic in $W \cap (M \times Q - N \times P)$. Hence, composing these homotopies, c is null-homotopic in $V \cap (M \times Q - N \times P)$. This completes the proof of (b) when $k = n - 2$.

Finally, if $k \leqq n - 3$ in case (b), we proceed as before to obtain K such that $P \subset \operatorname{Int} K$. Again by the special case of (b), $M \times Q - N \times K$ is 1-LC at each point of $N \times \operatorname{Int} K$ and one proceeds in the obvious manner to show that $M \times Q - N \times P$ is 1-LC at each point of $N \times P$. This completes the proof of case (b).

To consider case (c) we assume that $M^n - N^{n-1}$ is 1-LC at each point of N^{n-1}. We will show that $M^n \times Q^s - N^{n-1} \times P^{s-1}$ has 1-ALG at each point of $N \times P$. To this end let K be an s-manifold in $\operatorname{Int} Q$ such that $P \subset \operatorname{Int} K$. It follows from case (b) that $M \times Q - N \times K$ is 1-LC at each point of $N \times K$.

First consider a point x of $N \times P$ such that $x \in \operatorname{Int}(N \times K)$ and let T be any neighborhood, in $M \times Q$, of x. We suppose $T \cap \operatorname{Bd}(N \times K) = \emptyset$. We choose open sets U and V $(x \in V \subset U \subset T)$ so that:

i) Any loop in $V - (N \times K)$ is null homotopic in $U - (N \times K)$;

ii) $V \cap (N \times (K - P))$ has exactly two components which we denote by W_1 and W_2;

iii) The image under the inclusion induced homomorphism
$$i_*: H_1(V - N \times P) \to H_1(U - N \times P)$$
is infinite cyclic.

Now suppose $f: \operatorname{Bd} \Delta^2 \to V - (N \times P)$ represents a loop which bounds in $U - (N \times P)$. We wish to show that f can be extended to map Δ^2 into $U - (N \times P)$. By Lemma 5 we may assume that f is a homeomorphism and that $f(\operatorname{Bd} \Delta^2)$ is in special position with respect to $N \times K$. Now $f^{-1}(N \times K)$ consists of an even number, $2p$, of points. We proceed by induction on p. If $p = 0$ the desired result follows directly from i). In general, it can be shown, using iii) together with the fact that f bounds in $U - (N \times P)$, that there is an adjacent (on $\operatorname{Bd} \Delta^2$) pair of points of $f^{-1}(N \times K)$ which map into the same W_i, say W_1. We join these points by an arc A whose interior lies in $\operatorname{Int} \Delta^2$ and extend f to map A into W_1. We choose a neighborhood G of A in Δ^2 so that $\Delta^2 - G$ is the union of two disjoint disks D_1 and D_2. We extend f to map G into $V - (N \times P)$ in such a way that the arcs $\bar{G} \cap D_i$ $(i = 1, 2)$ are mapped into $U - (N \times K)$. One of the disks, say D_1, is such that $f(\operatorname{Bd} D_1) \subset V - (N \times K)$. By i) we can extend f to map D_1 into $U - (N \times K)$. Now $f \mid \operatorname{Bd} D_2$ bounds in $U - (N \times K)$ and $f(\operatorname{Bd} D_2) \cap (N \times K)$ consists of $2(p-1)$ points. By induction there is an extension of $f \mid \operatorname{Bd} D_2$ to a map of D_2 into $U - (N \times P)$. Combining these maps we obtain an extension of f which maps Δ^2 into $U - (N \times P)$, and the proof is complete.

In the case that $x \in N \times P$ is such that $x \in \operatorname{Bd}(N \times K)$ the proof is easier; for then we need only to choose arbitrarily small neighborhood pairs (U, V) satisfying i). The desired conclusion follows since any loop in $V - (N \times P)$ is homotopic in $V - (N \times P)$ to a loop in $V - (N \times K)$.

5. An example. Suppose that n is an integer such that there exists a pwl locally nice embedding $S^{n-2} \subset S^n$ with the two properties:

(a) $\pi_1(S^n - S^{n-2}) \approx Z$ (the infinite cyclic group);

(b) $\pi_k(S^n - S^{n-2}) \neq 0$ for some integer k, $2 \leqq k$, so that S^{n-2} is topologically knotted in S^n.

Such examples are known (see Neuwirth and Hirsch, *Piecewise-regular n-knots, Annals of Mathematics*, vol. 80 (1964), pp. 594-612). For instance, if one forms the pwl 10-manifold K^{10} by "plumbing" together two copies of the total space of the tangent bundle of S^5, then K^{10} can be embedded in S^{11} in a pwl locally flat manner and $\operatorname{Bd} K^{10}$ is the required S^9. The properties of this embedding can be verified using Neuwirth's construction of the universal covering space of $S^{11} - S^9$ and the fact that $S^{11} - K^{10}$ is 4-connected but not 5-connected. We obtain $\pi_5(S^{11} - S^9) \neq 0$. Neuwirth and Hirsch give such examples for $S^7 \subset S^9$ with $k = 4$ and it has been communicated to the authors that examples for $S^3 \subset S^5$ with $k = 2$ will be given in a forthcoming paper by C. T. C. Wall. For such a value of n and such an embedding $S^{n-2} \subset S^n$, we obtain:

PROPOSITION 2. *Let $\Sigma^j S^{n-2} = S^{n+j-2}$ denote the j-fold suspension of S^{n-2} in $\Sigma^j S^n = S^{n+j}$. Then, for $j \geqq 1$, the embedding $S^{n+j-2} \subset S^{n+j}$ is piecewise-linear and locally nice but not locally flat.*

Proof. Suppose first that $j = 1$. Subdivide S^n so that S^{n-2} is a full subcomplex, and let v be one of the suspension points for $\Sigma S^n = S^{n+1}$. Let B^{n+1} be the closed star of v in S^{n+1}. Since $B^{n+1} - B^{n+1} \cap S^{n-1}$ (where $S^{n-1} = \Sigma S^{n-2}$) has the homotopy type of $S^n - S^{n-2}$, which has fundamental group Z, $S^{n+1} - S^{n-1}$ has 1-ALG at v. Similarly for the other suspension point. At points of S^{n-1} other than the suspension points, the embedding is locally a product, so that by Theorem 2, $S^{n+1} - S^{n-1}$ has 1-ALG at each point of S^{n-1}.

Now, suppose that S^{n-1} were locally flat in S^{n+1} at v. Select an open set V such that the pair $(V, V \cap S^{n-1})$ is homeomorphic to the canonical pair (R^{n+1}, R^{n-1}) and so that $V \subset B^{n+1}$. Finally, subdivide S^{n+1} so that if F^{n+1} is the closed star of v in the subdivision, then $F^{n+1} \subset V$. Consider the inclusion homomorphisms:

$$\pi_k(F^{n+1} - S^{n-1}) \to \pi_k(V - S^{n-1}) \to \pi_k(B^{n+1} - S^{n-1}).$$

Now, $V - S^{n-1}$ has the homotopy type of S^1 and $k \geqq 2$, so the middle group is 0, while the composition of the two homomorphisms is an isomorphism. This is impossible, since the other two groups are not 0. Hence, S^{n-1} is not locally flat at v.

To continue the proof by induction on j, we need only check that

$\pi_1(S^{n+1}-S^{n-1})\approx Z$ and that $\pi_k(S^{n+1}-S^{n-1})\neq 0$. This follows from the fact that $S^{n+1}-S^{n-1}=\Sigma S^n-\Sigma S^{n-2}$ is homeomorphic to

$$(S^n-S^{n-2})\times(-1,1).$$

Combining Proposition 2 with the examples mentioned at the beginning of this section gives a locally nice but not locally flat pwl embedding of S^{n-2} in S^n for each $n\geqq 6$. It follows from Dehn's Lemma that such an example is impossible for $n=4$.

THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] P. Alexandroff, "On local properties of closed sets," *Annals of Mathematics* (2), vol. 36 (1935), pp. 1-35.

[2] M. Brown, "Locally flat imbeddings of topological manifolds," *Annals of Mathematics*, vol. 75 (1962), pp. 331-341.

[3] S. Eilenberg and R. L. Wilder, "Uniform local connectedness and contractibility," *American Journal of Mathematics*, vol. 64 (1942), pp. 613-622.

[4] O. G. Harrold, Jr., "Euclidean domains with uniformly abelian local fundamental groups," *Transactions of the American Mathematical Society*, vol. 67 (1949), pp. 120-129.

[5] ———, "Euclidean domains with uniformly abelian local fundamental groups, II," *Duke Mathematical Journal*, vol. 17 (1950), pp. 269-272.

[6] D. R. McMillan, Jr., "A criterion for cellularity in a manifold," *Annals of Mathematics*, vol. 79 (1964), pp. 327-337.

[7] J. R. Stallings, "The piecewise-linear structure of euclidean space," *Proceedings of the Cambridge Philosophical Society*, vol. 58 (1962), pp. 481-488.

[8] ———, "On topologically unknotted spheres," *Annals of Mathematics*, vol. 77 (1963), pp. 490-503.

AN EXTENSION OF RESULTS OF NOVIKOV AND BROWDER.

By C. T. C. Wall.

Let V^v be a closed, oriented, simply-connected, smooth manifold. According to Whitney [20], V^v can be smoothly imbedded in spheres S^{N+v} of sufficiently high dimension ($N \geqq v$ is certainly adequate). Let $z \in H_v(V)$ be the fundamental class of V, ξ^N the normal bundle of V in S^{N+v}, V^ξ the Thom complex of ξ, and $\alpha \in \pi_{N+v}(V^\xi)$ the element derived by the Thom construction ("shrinking the complement of an open tubular neighbourhood of V in S^{N+v} to a point"). The quadruple (V, z, ξ, α) has certain properties: e.g. (V, z) satisfies Poincaré duality, and the Hurewicz image of α, $h(\alpha)$, equals the image, $\phi(z)$, of z under the Thom isomorphism. In a remarkable paper [10], Novikov indicated a proof that in the case $N > v \geqq 5$, the invariants above almost characterised the diffeomorphism class of V; and Browder showed in [1] that most quadruples (V, z, ξ, α) satisfying the conditions could so arise.

The main object of this paper is to extend these results to the case of manifolds with boundary, where, somewhat surprisingly, a stronger result can be obtained. We also use a paper of Haefliger [2] (the methods of which were used in [5]) to extend the results to the metastable range: this extension was first investigated by Levine [7].

Since no proofs of any of the above results have appeared in print at the time of writing, we shall give them all here, making use only of the work of Milnor and Kervaire [6], [8].

1. Invariant systems. We first list the data which will be required for the construction. For a closed manifold, this was done above. It is convenient to have a name for them: we choose the name 'invariant systems,' but will always abbreviate it to I. S.; we distinguish the closed case from the bounded case by writing B. I. S. or C. I. S. We call the quadruple (X, z, ξ^N, α) a C. I. S. if X is a finite, simply-connected C. W. complex; $z \in H_v(X)$ has the property that for any integer r and abelian group G,

$$z \cap : H^r(X; G) \to H_{v-r}(X; G)$$

is an isomorphism (we call z a fundamental class for X); ξ^N is an N-vector

---

Received June 22, 1964.

bundle over X with Thom space X^ξ; and $\alpha \in \pi_{N+v}(X^\xi)$ has the property $h(\alpha) = \Phi(z)$.

The definition in the relative case is entirely analogous. We call $(X, \partial X, z, \xi, \alpha)$ a B. I. S. if $(X, \partial X)$ is a finite C. W.-pair, with each of X, ∂X connected and simply-connected; $z \in H_v(X, \partial X)$ induces isomorphisms

$$z \cap : H^r(X; G) \to H_{v-r}(X, \partial X; G)$$

and

$$z \cap : H^r(X, \partial X; G) \to H_{v-r}(X; G)$$

for all r, G; ξ^N is as above, and we write ∂X^ξ for the Thom complex of $\xi \mid \partial X$; and $\alpha \in \pi_{N+v}(X^\xi, \partial X^\xi)$ again satisfies $h(\alpha) = \Phi(z)$.

Since, by Whitney [20], any v-manifold has an essentially unique imbedding in D^{N+v} if $N > v$, we note that a smooth compact manifold determines an essentially unique B. I. S. or C. I. S., again provided $N > v$.

Lemma 1. *Let $(X, \partial X, z, \xi, \alpha)$ be a B. I. S. Then $(\partial X, \partial_* z, \xi \mid \partial X, \partial_* \alpha)$ is a C. I. S.*

Proof. By a standard property of cap products, the following diagram is commutative up to sign (with arbitrary coefficient group G):

$$\begin{tikzcd}
H^r(X, \partial X) \arrow[r] \arrow[d, "z \cap"] & H^r(X) \arrow[r] \arrow[d, "z \cap"] & H^r(\partial X) \arrow[r] \arrow[d, "\partial_* z \cap"] & H^{r+1}(X, \partial X) \arrow[r] \arrow[d, "z \cap"] & H^{r+1}(X) \arrow[d, "z \cap"] \\
H_{v-r}(X) \arrow[r] & H_{v-r}(X, \partial X) \arrow[r] & H_{v-r-1}(\partial X) \arrow[r] & H_{v-r-1}(X) \arrow[r] & H_{v-r-1}(X, \partial X).
\end{tikzcd}$$

Since both rows are exact, we can use the Five Lemma to conclude that

$$\partial_* z \cap : H^r(\partial X) \to H_{v-r-1}(\partial X)$$

is an isomorphism. It now only remains to observe

$$h(\partial_* \alpha) = \partial_* h(\alpha) = \partial_* \Phi(z) = \Phi(\partial_* z).$$

2. Constructing a manifold. What would be nice to prove is that there is a (1-1) correspondence between appropriately defined equivalence classes of B. I. S. and C. I. S. and embeddings of manifolds in spheres. This is not true, however, and one can only obtain approximations to it. We shall start by constructing an imbedded manifold which has the wrong homotopy type; this we shall try to rectify in subsequent paragraphs.

First, given a C. I. S., we can replace X by a homotopy equivalent simplicial complex, imbed this in Euclidean space of large dimension, and

201656

take the interior of a regular neighbourhood—this is a homotopy equivalent open differentiable manifold. For a B. I. S., we take $(X, \partial X)$ as a simplicial pair, and imbed in a simplex Δ^k in such a way that ∂X is imbedded in one face, and the remainder of X in the interior. Taking a neighbourhood in Δ^k now allows us to replace X by a (non-compact) manifold, ∂X by its boundary.

We can now regard the Thom complex X^ξ as a manifold (except at the 'point at infinity') with boundary ∂X^ξ. In the case of a C. I. S. we now take a representative $f: S^{N+v} \to X^\xi$ of α, make it t-regular on X (c.f. Thom [14]), and write $M^v = f^{-1}(X)$; this is a smooth submanifold of S^{N+v}, with normal bundle $f^*(\xi)$ (by the t-regularity). For a B. I. S., we represent α by $f: (D^{N+v}, \partial D^{N+v}) \to (X^\xi, \partial X^\xi)$, and make this t-regular on $(X, \partial X)$. Observe particularly that if a map from ∂D^{N+v} to ∂X^ξ, already t-regular on ∂X, is in the homotopy class $\partial_* \alpha$, we can take this as the restriction of f to ∂X, by another result of Thom (loc. cit.). Again, when f is t-regular on X, $M^v = f^{-1}(X)$ is a smooth submanifold of D^{N+v}, meeting the boundary transversely in $\partial M = f^{-1}(\partial X)$, and with normal bundle in D^{N+v} given by $f^*(\xi)$.

The construction, which is due to Browder [1], gives us a manifold M^v, and a map $f: M \to X$ with $f^{-1}(\partial X) = \partial M$. We assert that f has degree 1. For let $\eta = f^* \xi$ be the normal bundle of M. Then we have maps $S^{N+v} \to M^\eta \to X^\xi$ (in the closed case) or

$$(D^{N+v}, \partial D^{N+v}) \to (M^\eta, \partial M^\eta) \to (X^\xi, \partial X^\xi);$$

the first has degree 1 by construction, and the composite also, by the assumption $h(\alpha) = \Phi(z)$. Hence the map from M^η to X^ξ has degree 1; desuspending by the Thom isomorphism, so has f.

LEMMA 2. *Let M, X satisfy Poincaré duality with fundamental classes y, z. Let $f: M \to X$ have degree* 1. *Write $K^r(M; G)$ for the cokernel of $f^*: H^r(M; G) \leftarrow H^r(X; G)$. Then f_* induces a map from $K^r(M; G)$ to $H^r(M; G)$ so that*

$$H^r(M; G) \cong K^r(M; G) \oplus f^* H^r(X; G).$$

f^* is a monomorphism of rings, and

$$f^* H^r(X; G) \cup K^s(M; H) \subset K^{r+s}(M; G \otimes H).$$

An analogous splitting occurs in homology.

Proof. Since $f_*(y) = z$, the diagram

$$\begin{tikzcd}
H^r(M;G) \arrow[d, "y\cap"] & H^r(X;G) \arrow[l, "f^*"'] \arrow[d, "z\cap"] \\
H_{v-r}(M;G) \arrow[r, "f_*"] & H_{v-r}(X;G)
\end{tikzcd}$$

is commutative, for by a standard property of cap products,

$$z\cap c = f_*y \cap c = f_*(y\cap f^*c).$$

Since $z\cap$ is an isomorphism, $(z\cap)^{-1}f_*(y\cap)$ splits f^*. We have $K^r(M;G) = (y\cap)^{-1}\operatorname{Ker} f_*$. If $c_1 \in H^r(X;G)$, and $c_2 \in K^s(M;H)$, then

$$\begin{aligned} f_*(y\cap c_2\cdot f^*c_1) &= f_*((y\cap c_2)\cap f^*c_1) \\ &= f_*(y\cap c_2)\cap c_1 = 0\cap c_1 = 0, \end{aligned}$$

so $c_2\cdot f^*c_1$ is in $K^{r+s}(M;G\otimes H)$.

We shall also need the corresponding results for bounded manifolds. Let $y\in H_v(M,\partial M)$ and $z\in H_v(X,\partial X)$ be fundamental classes; $f:(M,\partial M)\to(X,\partial X)$ have degree 1 (i. e. $f_*y=z$); $K_r(M) = \operatorname{Ker} f_*: H_r(M)\to H_r(X)$; similarly for $K_r(\partial M)$ and $K_r(M,\partial M)$.

LEMMA 3. *With the above assumptions, the exact homology sequence of* $(M,\partial M)$ *splits as the direct sum of an isomorphic copy of the exact sequence of* $(X,\partial X)$ *with the sequence of the* K_r. *The intersection of any* $x\in K_r(M;G)$ *with any* $y\in H_{v-r}(X,\partial X;H)$ *vanishes.*

Proof. Consider the diagram

$$\begin{tikzcd}
H^{v-r}(X,\partial X;G) \arrow[r] \arrow[d, "f^*"] & H^{v-r}(X;G) \arrow[r] \arrow[d, "f^*"] & H^{v-r}(\partial X;G) \arrow[r] \arrow[d, "f^*"] & H^{v-r+1}(X,\partial X;G) \arrow[d, "f^*"] \\
H^{v-r}(M,\partial M;G) \arrow[r] \arrow[d, "y\cap"] & H^{v-r}(M;G) \arrow[r] \arrow[d, "y\cap"] & H^{v-r}(\partial M;G) \arrow[r] \arrow[d, "\partial_*y\cap"] & H^{v-r+1}(M,\partial M;G) \arrow[d, "y\cap"] \\
H_r(M;G) \arrow[r] \arrow[d, "f_*"] & H_r(M,\partial M;G) \arrow[r] \arrow[d, "f_*"] & H_{r-1}(\partial M;G) \arrow[r] \arrow[d, "f_*"] & H_{r-1}(M;G) \arrow[d, "f_*"] \\
H_r(X;G) \arrow[r] & H_r(X,\partial X;G) \arrow[r] & H_{r-1}(\partial X;G) \arrow[r] & H_{r-1}(X;G)
\end{tikzcd}$$

As in Lemma 2, the composite map of the upper row on the lower is the isomorphism induced by $z\cap$. Since the maps induced by $y\cap$ are isomorphisms, f^* again splits f_* additively. By the commutativity of the lower squares, the K's form a subcomplex of the exact sequence of $(M, \partial M)$; since $(y \cap) f^* (z \cap)^{-1}$ also imbeds the exact sequence of $(X, \partial X)$ as a subcomplex, and the groups in the third row all split, the whole sequence splits as a direct sum.

For the last part, recall that intersections can be computed as cup products; we wish to show

$$K^{v-r}(M, \partial M; G) \cdot f^* H^r(X; H) = 0.$$

But by an argument as in Lemma 2, we see that the product lies in $K^v(M, \partial M; G \otimes H) = 0$.

We remark that the sequence of K's looks rather like the sequence of homology groups of a manifold, except that the extreme terms are absent. In particular, the direct sum splitting shows that the K's satisfy Poincaré or Lefschetz duality (appropriately interpreted) and the universal coefficient theorem.

3. Surgery. In § 2 we constructed from a C. I. S. or B. I. S. a manifold M^v embedded in S^{N+v} or D^{N+v} and a map of degree 1 $f: M^v \to X^v$ such that $f^*\xi$ is isomorphic to the normal bundle η of M, and if the map from S^{N+v} resp. D^{N+v} to M^η given by the Thom construction is followed by the map from M^η to X^ξ determined by an isomorphism above, the composite has homotopy class α. Under these circumstances we say M *corresponds* to the C. I. S. or B. I. S. If also f is a homotopy equivalence, we say M *realises* the invariant system.

Let M correspond to an I. S.; we shall simplify M by surgery, making sure that at each stage it corresponds to the I. S., and try to make f a homotopy equivalence. We shall find it necessary to impose arithmetic restrictions on v and N as we proceed.

Suppose inductively f r-connected $(r \geqq 0)$, and $x \in K_r(M; \mathbf{Z})$.

(3.1) *If $r > 0$, x is spherical.*

For $r = 0$, $K_0(M) = \tilde{H}_0(M)$ is generated by spherical classes, though it does not consist solely of them. If $r = 1$, M is connected, and $x \in K_1(M) = H_1(M)$ is always spherical. If $r \geqq 2$, M and X are simply-connected, and by the relative Hurewicz theorem, $h: \pi_{r+1}(f) \to H_{r+1}(f)$ is an isomorphism. But ∂_* induces an isomorphism of $H_{r+1}(f)$ on $K_r(M)$, so we can write $x = \partial_* h(y) = h \partial_*(y)$, and x is spherical.

(3.2) *If $v \geqq 2r + 1$, x can be represented by an imbedded sphere S^r.*

By (3.1), we can represent x by a map of S^r to M, and by a general position argument which stems from Whitney [20], we can deform such a map slightly to an imbedding, i say. We can also obtain an imbedding if $v = 2r$, $r > 2$ and M is simply-connected [3].

We now use $i(S^r)$ to perform surgery to kill x; however, we must be careful that our construction gives an M' which corresponds to the given I. S. We shall in fact construct a submanifold W^{v+1} of $S^{N+v} \times I$ or $D^{N+v} \times I$ and a homotopy $F\colon S^{N+v} \times I \to X^{\xi}$ or $(D^{N+v}, \partial D^{N+v}) \times I \to (X^{\xi}, \partial X^{\xi})$ which is t-regular on X, has $F^{-1}(X) = W$, and $F(P, 0) = f(P)$ for $P \in S^{N+v}$ or D^{N+v}. Then, clearly, if M' is the 'upper end' $W \cap (S^{N+v} \times 1)$, M' corresponds to the I. S., and if W is diffeomorphic to $M \times I$ with an $(r+1)$-handle attached to kill $x \in K_r(M)$, we are one step forward in the surgery.

(3.3) *If* $v \geqq 2r$ *and* $N > r+1$, *we can perform surgery as above to kill* $x \in K_r(M)$ (*if* $r \geqq 2$), $x \in \pi_r(M)$ (*if* $r = 0, 1$), *provided if* $v = 2r$ *that* $r \geqq 4$, *and* $r \neq 7$ *and the sphere representing* x *has trivial normal bundle in* M.

Since $fi\colon S^r \to X$ is nullhomotopic, we can extend to a map j of D^{r+1}. The bundle $j^*\xi$ is trivial: we choose a framing $f_1, \cdots, f_N$ for it. As $j^*\xi \mid S^{r-1} = i^*f^*\xi = i^*\eta$, we have now a framing along $i(S^r)$ of the normal bundle η of M in S^{N+v}. Deform $i(S^r)$ along the normal vector f_1 into the complement of M, and then span by a disc D^{r+1}, which we can suppose smooth. Since $N + v \geqq 3r + 2 \geqq 2r + 3$, we may suppose D^{r+1} imbedded. As also $N + v > (r+1) + v$, we can suppose D^{r+1} disjoint from M except along the boundary.

The vectors $f_2, \cdots, f_N$ give a normal $(N-1)$-field to D^{r+1} along its boundary. Since the normal bundle of D^{r+1} is trivial, we can regard them as defining a map from S^r to the Stiefel manifold $V_{N+v-r-1, N-1}$. If $r < v - r$, all such maps are trivial, so we can extend $f_2, \cdots, f_N$ to a framing $f_2, \cdots, f_{N+v-r}$ over all D^{r+1}. If $r = v - r$, and $r \neq 1, 3, 7$ since

$$\partial_*\colon \pi_r(V_{N+v-r-1, N-1}) \to \pi_{r-1}(SO_{v-r})$$

is then a monomorphism, and the image of the element above is clearly the characteristic class of the normal bundle of $i(S^r)$, which vanishes by hypothesis, we can again extend. Use the vectors $f_{N+1}, \cdots, f_{N+v-r}$ to define a tube $D^{r+1} \times D^{v-r}$ round D^{r+1}, meeting M in $S^r \times D^{v-r}$.

Now take $M \times I \subset S^{N+v} \times I$, and attach $D^{r+1} \times D^{v-r} \times 1$ in $S^{N+v} \times 1$. There is a right-angle corner at $S^r \times S^{v-r}$; this is to be straightened, and $D^{r+1} \times \mathrm{Int}(D^{v-r})$ pushed into $S^{N+v} \times (0, 1)$. For more details of this construction, see Haefliger [2], p. 460. This defines W. We now define the homotopy F. We start with $f\colon M \to X$, the given map extended (as given) over a tubular neighbourhood of M to X^{ξ} and (we may suppose) mapping the remainder of S^{N+v} to the point at ∞. Extend over W by mapping $M \times I$ via projection on M, D^{r+1} by j, and the remainder of $D^{r+1} \times D^{v-r}$ using a retraction on $(D^{r+1} \times 0) \cup (S^r \times D^{n-r})$, which is already mapped. Over a tubular neighbourhood of W, for $M \times I$ we again use projection; over D^{r+1} the normal framing $f_0, f_2, \cdots, f_N$ (where f_0 is the direction of the second factor in $S^{N+v} \times I$) defines a map (recall that $f_1, \cdots, f_N$ framed $j^*\xi$). When the corner is straightened, f_0 and f_1 will be forced to fit together there; we then have covered $h: W \to X$ by a map in ξ of the normal bundle of W in $S^{N+v} \times I$. Take the induced map of a tubular neighbourhood of W to the Thom complex of ξ, and map the rest of $S^{N+v} \times I$ to the point ∞. The constructed map has all desired properties and we have completed the proof of (3.3).

(3.4) *Let* $v = 2k$ *or* $2k+1$, $N > k \geqq 2$. *Then we can perform surgery to make* M *simply-connected, and* $f: M \to X$ k*-connected, so that* $K_r(M) = 0$ *for* $r < k$.

We first add 1-handles to connect up the components of M. Next kill in turn the generators of the fundamental group of M; under the assumption $k \geqq 2$ above, it follows from ([6] Lemma 5.2) that this reduces the group to zero. The argument above will now permit us to kill in turn the $K_i(M)$ for $2 \leqq i < k$.

In the case when M is bounded, we can perform surgery both on ∂M and on M. By an observation made in § 2, when V is given, and $f: S^{N+v-1} \to \partial X^{\xi}$ represents α, is t-regular on ∂X, and has $V = f^{-1}(\partial X)$, we can take M as above with boundary V, and then perform surgery on M.

(3.5) *Let* $v = 2k+1$, $N > k \geqq 2$. *Then we can perform surgery to make* M *and* ∂M *simply-connected, and kill all the* K_r *except those lying in the exact sequence*

$$0 \to K_{k+1}(M) \to K_{k+1}(M, \partial M) \to K_k(\partial M) \to K_k(M) \to K_k(M, \partial M) \to 0.$$

We may first perform surgery on ∂M to make $K_r(\partial M) = 0$ for $r < k$, then the same for M. The higher $K_i(\partial M)$ and $K_i(M, \partial M)$ vanish by duality (Lemma 3) and the exact sequence shows that the other K's are also zero.

(3.6) *Let* $v = 2k+2$, $N - 1 > k \geqq 2$. *Then we can perform surgery to make* M *and* ∂M *simply-connected and kill all the* K_r *except*

$$0 \to K_{k+1}(\partial M) \to K_{k+1}(M) \to K_{k+1}(M, \partial M) \to K_k(\partial M) \to 0.$$

The proof is as for (3.5).

4. Surgery in middle dimensions.

(4.1) *Suppose M closed, $v=2k+1$, $N-1>k\geqq 2$. Then we can perform surgery to kill K_k also and so make $f: M\rightarrow X$ a homotopy equivalence.*

By (3.3), the above hypotheses are sufficient to ensure that for any $x\in K_k(M)$, we can perform surgery to kill it. Now by § 5 of [6] (particularly Lemma 5.7) or § 2 of [15] the result follows if k is even. When k is odd, the argument of § 6 of [6] gives the result[1] provided that (3.3) can be sharpened to state that the normal framing of $i(S^r)$ in M induced by the modification can be altered by any element in the image of

$$\partial: \pi_{k+1}(S^{k+1})\rightarrow \pi_k(SO_{k+1}).$$

But this is not difficult (c.f. [2], end of § 3). For the extension of the partial frame $f_2,\cdots,f_N$ over D^{r+1} may be varied by any element α of $\pi_{r+1}(V_{N+v-r-1,N-1})$ (it is always possible to extend the resulting partial frame over the contractible D^{r+1} to a complete frame). Substituting for v and r in terms of k, this group is $\pi_{k+1}(V_{N+k,N-1})$. Now the Stiefel manifold is the homogeneous space SO_{N+k}/SO_{k+1}, hence the above lies in the exact sequence

$$\pi_{k+1}(V_{N+k,N-1})\xrightarrow{\partial_*}\pi_k(SO_{k+1})\xrightarrow{i_*}\pi_k(SO_{N+k}).$$

The varying by α of the partial frame over D^{r+1} alters the normal framing of $i(S^r)$ in M by $\partial_*\alpha$; this can be any element of $\operatorname{Im}\partial_*=\operatorname{Ker} i_*$. But as $N\geqq 2$, $\pi_k(SO_{N+k})$ is stable; $\operatorname{Ker} i_*$ is the same as for $N=2$, hence equals the image of $\pi_{k+1}(V_{k+2,1})=\pi_{k+1}(S^{k+1})$.

(4.2) *Suppose M bounded, $f\mid\partial M$ already a homotopy equivalence, and f k-connected ($v=2k+1\geqq 5$, $N>k+1$). Then we can perform surgery to make f a homotopy equivalence.*

By a remark at the end of § 3, the K_r satisfy duality and the universal coefficient theorem. Hence in the middle dimension, they have linking numbers with the usual properties. The argument of Kervaire and Milnor is now again valid.

We can now give

THEOREM 1A. (Browder [1]). *Let $v>4$, $v\neq 2$ (mod 4), $2N>v+2$.*

---

[1] It is possible to show that the argument on p. 525 of [6], using Lemma 5.10, can be applied here, by obtaining Wu formulae for X. However, this argument of [6] is irrelevant. For since L is nonsingular, to λ corresponds a μ with $L(\lambda,\mu)\neq 0$, and hence the class in M' determined by μ has order $\neq 2$. This rules out the second possibility on p. 526, the only one which causes trouble.

Then a C.I.S. can be realised by a manifold, provided, if $v = 4j$, *that* $\sigma = \tau$, *where* σ *is the signature of* X *and* τ *the (Hirzebruch) L-genus of the bundle inverse to* ξ.

Proof. Only the case $v = 4j$ remains, and we can mimic the proof in [6] provided that the signature of K_{2j} vanishes. But by Hirzebruch's theorem, τ is the signature of M, so this condition is equivalent to $\sigma = \tau$.

For a more detailed discussion of the case $v = 4j + 2$, we refer the reader to work of S. P. Novikov [11].

COMPLEMENT 1B. *If* X *is a closed smooth manifold, the result holds also if* $v = 4$.

The above proof breaks down, as we do not have an adequate imbedding theorem for S^2 in M^4. However, we know that M and X have the same signature, and that $w_2(M)$ and $w_2(X)$ are both zero or both nonzero. By (4.2) of [18], if we take the connected sum of M with sufficiently many copies of $S^2 \times S^2$, the result is diffeomorphic to a connected sum of X with copies of $S^2 \times S^2$. This can be achieved by performing spherical modifications on circles which bound discs in M. Now using Theorem 2 of [19], we may suppose that M is the connected sum of X with $k \geqq 2$ copies of $S^2 \times S^2$, and that the map of M to X is induced by shrinking these to a point. The remainder of the proof now proceeds as usual.

Our main result is

THEOREM 2. *Let* $v \geqq 5$, $2N > v + 2$. *Then any B.I.S. can be realised by a manifold (provided if* $v = 5$ *that* ∂X *is a closed smooth manifold)* $V^v \subset D^{N+v}$. *If* $v \geqq 6$, $2N > v + 3$, *the pair* (D^{N+v}, V^v) *is unique up to diffeomorphism.*

Proof. First suppose $v = 2k + 1$, $k \geqq 2$, $N > k + 1$. Then it follows by (4.2) that any manifold ∂V realising the boundary of the B.I.S. is the boundary of a manifold V realising the B.I.S. So we must check existence of ∂V. If k is even, we can observe that for ∂X, $\sigma = \tau = 0$; the proof of this is identical with the proof (Thom [13]) valid for the boundary of a smooth manifold. The result then follows from 1A and 1B. In fact we can give a geometrical argument which covers also the case when k is odd.

Use (3.5) to kill all the K_i except those in the middle dimensions. Next, we will kill $K_k(M, \partial M)$ (c.f. [18], Theorem 1 for the proof). By (4.2) we may suppose that $K_k(\partial M) \neq 0$, so is free abelian. Since the homomorphisms $K_{k+1}(M, \partial M) \rightarrow K_k(\partial M)$ and $K_k(\partial M) \rightarrow K_k(M)$ are dual, $K_k(M)$ has nonzero rank. So it is generated by primitive (i.e. indivisible) elements.

Now if $K_k(M, \partial M)$ is nonzero, we choose $x \in K_k(M)$ primitive and with nonzero image in $K_k(M, \partial M)$, and perform surgery starting with x. Since x is primitive, no new element is introduced in $K_k(M, \partial M)$, the effect is simply to kill the image of x. Each such step decreases either the rank or the order of the torsion subgroup of $K_k(M, \partial M)$; making an induction on these, we see that we can reduce the group to zero.

By duality, $K_{k+1}(M) = 0$, so only the terms of

$$0 \rightarrow K_{k+1}(M, \partial M) \xrightarrow{b} K_k(\partial M) \rightarrow K_k(M) \rightarrow 0$$

survive.

Now these groups are all free, so the extension splits, and it follows (as in [6] Lemma 7.1), that if we can kill in turn the elements of a basis of $\operatorname{Im} b$, we reduce $K_k(\partial M)$ to zero, and the result follows. To prove these elements spherical, we need a 'relative relative' Hurewicz theorem for the quadruple

$$\Phi\colon \begin{tikzcd} \partial M \arrow[r] \arrow[d] & M \arrow[d] \\ \partial X \arrow[r] & X. \end{tikzcd}$$

Since all four spaces are 1-connected (hence the map $\partial X \rightarrow X$ also is), and the map $M \rightarrow X$ is k-connected, we can apply Theorem 2.4 of [9] with $m = 2$, $n = k + 1$, $r = k + 1$. For $q \leqq k + 1$, $H_q(\Phi) = K_{q-1}(M, \partial M) = 0$, so Φ is $(k+1)$-connected, and $\pi_{k+2}(\Phi)$ maps onto $H_{k+2}(\Phi) \cong K_{k+1}(M, \partial M)$. So the (isomorphic) image of $K_{k+1}(M, \partial M)$ in $H_{k+1}(M, \partial M)$ consists of spherical classes.

Take any element of $\operatorname{Im} b$, and lift it to an element of $K_{k+1}(M, \partial M)$. By the above, we may represent this by a map of a disc $f\colon D^{k+1} \rightarrow M$ with $f(\partial D^{k+1}) \subset \partial M$, and we may suppose the boundary k-sphere imbedded. Take the map of the disc in general position (c.f. [21]): then the singularity set Σ of f is 1-dimensional, and $f(\Sigma)$ is a union of arcs and simple loops. The map f is an immersion except at a finite set of points corresponding to the end-points of these arcs. We assert that f is homotopic $(\operatorname{rel} \partial D^{k+1})$ to an immersion. For by [4], the only obstruction to this lies in $H^{k+1}(D^{k+1}, \partial D^{k+1}; \pi_k(V_{2k+1,k+1})) \cong \pi_k(V_{2k+1,k+1})$, a group isomorphic either to $\mathbf{Z}$ or to $\mathbf{Z}_2$. But each singularity point (being generic), contributes a unit to the obstruction: since there are an even number of them (each arc of $f(\Sigma)$ contains just 2), the obstruction is zero in the $\mathbf{Z}_2$ case: it vanishes also in the integer case, as it is not difficult to verify that the singularities at the 2 ends of an arc have opposite signs. Thus we may suppose f an immersion:

then $f(D^{k+1})$ has trivial normal bundle, and $f(\partial D^{k+1})$ also has. Thus by (3.3) we can perform surgery to kill elements of $\operatorname{Im} b$, and the result follows. The appeal to (3.3) fails if $r = 3, 7$. However, the existence of the immersed disc shows directly that the obstruction there obtained in $\pi_k(V_{N+k-1,N-1})$ vanishes. For this obstruction is identified (by suspension) with the obstruction in $\pi_k(V_{2k+1,k+1})$ to immersing the disc. Geometrically, if the disc is immersed in M^v, $\tau_M \mid D^{k+1}$ gives a $(2k+1)$-frame, and the normal vectors a complementary N-frame. So the isomorphic image of the obstruction in $\pi_k(V_{N+k,N})$ vanishes.

Now we let $v = 2k$, $k \geqq 3$, $N > k$. By (4.1), there is a manifold ∂M realising the boundary C.I.S.; this bounds some M, and by (3.6) we can kill all K's except $K_k(M) \cong K_k(M, \partial M) \cong K^k(M)$. If k is even, this is a free abelian group, with nonsingular even quadratic form (c.f. [8], Lemma 7 and Remark), hence has signature divisible by 8. If k is odd, then by the method of [6, § 8] we can define an Arf invariant c of $K_k(M)$. According to [16], there exists an almost-closed $(k-1)$-connected π-manifold H of dimension $2k$, with signature minus that of $K_k(M)$ when k is even; with Arf invariant c when k is odd, moreover, H is a handlebody. $H \times I$, with corners rounded, is then also a handlebody, so by the classification ([17] Theorem 2) is diffeomorphic to a sum of copies of $S^k \times D^{k+1}$. So it imbeds in $\boldsymbol{R}^{2k+1}$, and H imbeds in this, so also in S^{N+2k-1}, with trivial normal bundle (this result is due to Milnor). Take such an imbedding, and deform the interior of H into that of D^{N+2k}. Now form the (boundary-connected) sum of (D^{N+2k}, M) with (D^{N+2k}, H) to obtain a pair (D^{N+2k}, N) with normal bundle η'. Since the normal bundle of H in D^{N+2k} is trivial, we can extend $f\colon M^\eta \to X^\xi$ to a map of $N^{\eta'}$, using the framing of H to map the second part to S^N, the Thom space of the base point $*$ in ∂X. Clearly the new map of D^{N+2k} is still in the homotopy class $\alpha \in \pi_{N+2k}(X^\xi, \partial X^\xi)$, and it is t-regular on X, with inverse image N. However, as H is mapped to $*$, $K_k(N) = K_k(M) \oplus H_k(H)$, so by construction it has signature (resp. Arf invariant) zero. The proof is concluded as before by appealing to Lemma 7.3 or 8.4 of [6].

Now we consider uniqueness. Let $(D^{N+v}, V_1{}^v)$ and $(D^{N+v}, V_2{}^v)$ be two pairs, realising the given I.S. with maps $f_i\colon (D^{N+v}, \partial D^{N+v}) \to (X^\xi, \partial X^\xi)$. Take a homotopy F of f_1 to f_2, and make it also t-regular on X (keeping the ends fixed), so that $F^{-1}(X) = W^{v+1}$ is a smooth manifold with corner, and boundary in 3 parts—V_1, V_2, and $\partial_c W = W \cap (\partial D^{N+v} \times I)$. We now perform surgery on W as above.

First consider a C.I.S. If $v = 2k$, $N - 1 > k \geqq 2$, then as in (4.1) we can perform surgery to make W, too, map by a homotopy equivalence. We then have an h-cobordism of pairs (S^{N+v}, V^v); by a result of Smale [12], since $v \geqq 5$, $N \geqq 3$, this is diffeomorphic to a product, so the pairs are diffeomorphic. If $v = 2k+1$, there is the usual obstruction to performing the surgery, and we find

THEOREM 1C. (Novikov [10]). *With the assumptions of* 1A, *and if* $2N > v+3$, *the corresponding manifold is unique up to diffeomorphism if* v *is even, and to connected sum with a pair* (S^{N+v}, Σ^v) (*where* Σ^v *bounds a* π-*manifold*) *if* v *is odd.*

We return to the uniqueness part of Theorem 2. If v is odd we perform surgery on $\partial_c W$ to obtain an h-cobordism as above. If v is even, we again avoid the obstruction by taking a boundary-connected sum with a pair (D^{N+v+1}, H), as in the existence part of the proof. So we may suppose $\partial_c W$ an h-cobordism. Now if v is even, we can perform surgery on W (leaving ∂W fixed) to obtain a homotopy equivalence. If v is odd, there appears to be an obstruction, but the same proof as above shows that it vanishes, so we can make F a homotopy equivalence then also. We now apply Smale's theorem twice, first to see that $(S^{N+v-1} \times I, \partial_c W)$ is diffeomorphic to a product, so that a neighbourhood of it also is, and now again to $W - \partial_c W$, using the product structure just obtained in the complement of a compact set, to say that it extends over W. Thus $(D^{N+v} \times I, W)$ is diffeomorphic to the product $(D^{N+v}, V_1{}^v) \times I$, and the two pairs $(D^{N+v}, V_1{}^v)$ and $(D^{N+v}, V_2{}^v)$ are diffeomorphic, as asserted.

TRINITY COLLEGE,
CAMBRIDGE.

Now at: DEPARTMENT OF PURE MATHEMATICS,
THE UNIVERSITY, LIVERPOOL 3.

REFERENCES.

[1] W. Browder, "Homotopy type of differentiable manifolds," *Colloquium on algebraic topology notes*, Aarhus, 1962, pp. 42-46.

[2] A. Haefliger, "Knotted $(4k-1)$-spheres in $6k$-spaces," *Annals of Mathematics*, vol. 75 (1962), pp. 452-466.

[3] ———, "Plongements différentiables de variétés dans variétés," *Commentarii Mathematici Helvetici*, vol. 36 (1961), pp. 47-82.

[4] M. W. Hirsch, "Immersions of manifolds," *Transactions of the American Mathematical Society*, vol. 93 (1959), pp. 242-276.

[5] W. C. Hsiang, J. Levine and R. H. Szczarba, "On the normal bundle of a homo-

topy sphere embedded in euclidean space," *Topology*, vol. 3 (1965), pp. 173-182.

[6] M. A. Kervaire and J. W. Milnor, "Groups of homotopy spheres I," *Annals of Mathematics*, vol. 77 (1963), pp. 504-537.

[7] J. Levine, "On differentiable imbeddings of simply-connected manifolds," *Bulletin of the American Mathematical Society*, vol. 69 (1963), pp. 806-809.

[8] J. W. Milnor, "A procedure for killing homotopy groups of differentiable manifolds," *Proceedings of the American Mathematical Society Symposium in Pure Mathematics*, III, 1961.

[9] I. Namioka, "Maps of pairs in homotopy theory," *Proceedings of the London Mathematical Society*, 3rd Ser. 12 (1962), pp. 725-738.

[10] S. P. Novikov, "Diffeomorphisms of simply-connected manifolds," *Soviet Mathematics Doklady*, vol. 3 (1962), pp. 540-543.

[11] ———, "Some properties of $(4k+2)$-dimensional manifolds," *Soviet Mathematics Doklady*, vol. 4 (1963), pp. 1768-1772.

[12] S. Smale, "On the structure of manifolds," *American Journal of Mathematics*, vol. 84 (1962), pp. 387-399.

[13] R. Thom, "Espaces fibrés en sphères et carrés de Steenrod," *Annales Scientifiques de l'Ecole normale supérieure*, vol. 69 (1952), pp. 109-182.

[14] ———, "Quelques propriétés globales des variétés différentiables," *Commentarii Mathematici Helvetici*, vol. 28 (1954), pp. 17-86.

[15] C. T. C. Wall, "Killing the middle homotopy groups of odd dimensional manifolds," *Transactions of the American Mathematical Society*, vol. 103 (1962), pp. 421-433.

[16] ———, "Classification of $(n-1)$-connected $2n$-manifolds," *Annals of Mathematics*, vol. 75 (1962), pp. 163-189.

[17] ———, "Classification problems in differential topology I: Classification of handlebodies," *Topology*, vol. 2 (1963), pp. 253-261.

[18] ———, "On simply-connected 4-manifolds," *Journal of the London Mathematical Mathematical Society*, vol. 39 (1964), pp. 141-149.

[19] ———, "Diffeomorphisms of 4-manifolds," *Journal of the London Mathematical Society*, vol. 39 (1964), pp. 131-140.

[20] H. Whitney, "Differentiable manifolds," *Annals of Mathematics*, vol. 37 (1936), pp. 645-680.

[21] ———, "The singularities of a smooth n-manifold in $(2n-1)$-space," *Annals of Mathematics*, vol. 45 (1944), pp. 247-293.

DERIVATIONS AND AUTOMORPHISMS OF COMPLETE REGULAR LOCAL RINGS.[1]

By NICKOLAS HEEREMA.

I. Introduction. Let R be a complete regular local ring with maximal ideal M and residue field k. In the unequal characteristic case we assume R unramified, that is, if p is the characteristic of k, then p is not in M^2. We reserve the symbol ϵ for the identity map on a ring which may in general be identified by the context. For r a positive integer, let G_r denote the group of automorphisms α of R such that $\alpha - \epsilon$ maps R into M^r. Let H_r be the subgroup of those α in G_r with the property that $\alpha - \epsilon$ maps M into M^{r+1}.. We have then the sequence of ramification groups.

$$G_1 \supset H_1 \supset G_2 \supset \cdots \supset G_n \supset H_n \supset \cdots \tag{1.1}$$

each group of which is a normal subgroup of the automorphism group of R. The elements of G_1 are called inertial automorphisms.

If R is a complete discrete valuation ring then, in the unequal characteristic case, $G_n = H_n$ for all n and the factor groups G_i/G_{i+1} were found by MacLane [6, Theorem 15] to be isomorphic to the direct product of the additive group of k to as many factors as the number of elements in a p-basis for k. This last group is the same as the additive group of derivations of k into k. In Section II of this paper we determine the factor groups G_i/H_i and H_i/G_{i+1} for $i = 1, 2, \cdots$ in the case in which R is an arbitrary unramified complete regular local ring. These results are described in Theorems 2.1, 2.2, and 2.3.

Let $\mathfrak{J}$ denote the set of all $n+1$ tuples of non-negative integers including $Q = (0, \cdots, 0)$. Let T be a ring. A higher derivation on $n+1$ indices is a sequence of maps $\{D^{(I)}\}_{I \in \mathfrak{J}}$ of T into T such that $D_Q = \epsilon$ and, for all $I \in \mathfrak{J}$,

$$D^{(I)}(a+b) = D^{(I)}(a) + D^{(I)}(b),$$

and

$$D^{(I)}(ab) = \sum_{Q \leqq J \leqq I} D^{(I-J)}(a) D^{(J)}(b). \tag{1.3}$$

The statement $J \leqq I$ means that each component of J is less than or equal

---

Revised received May 28, 1965.

[1] This research was supported by NSF GP-1084.

to the corresponding component of I and $I-J$ is obtained by component wise subtraction.

We consider a complete local ring R with maximal ideal $M=(X_0,\cdots,X_n)$. Let $D=\{D^{(I)}\}$ be a higher derivation on R and let $X^I=X_0^{i_0}\cdots X_n^{i_n}$ where $I=(i_0,\cdots,i_n)$. The map α_D on R defined by

$$\alpha_D(A)=\sum_{I\geqq Q}D^{(I)}(A)X^I \tag{1.4}$$

is an inertial automorphism provided the determinant of D (see Section III) is different from zero. Theorem 3.1 states that every α in G_1 has the form (1.4) if R is an unramified regular complete local ring and $\{X_0,\cdots,X_n\}$ is a minimal basis for M. In the case in which R is a power series ring over a field of characteristic zero or a complete discrete valuation ring of characteristic zero, the representation (1.4) leads to a Taylor series like representation of inertial automorphisms using derivations. (Theorems 3.2 and 3.3).

In Section IV we define a composition $*$ on $\mathfrak{H}$, the set of all higher derivations on R, which is preserved under the map $D\rightarrow\alpha_D$. The final result, Theorem 4.2 asserts that, in the equal characteristic case, a suitably defined subset $\mathfrak{H}_k$ of $\mathfrak{H}$ is a group with respect to $*$, and is isomorphic to G_1.

II. Factors of the ramification groups. It is well known, and was proved by I. S. Cohen [1, Theorem 15] that a complete regular local ring R of dimension $n+1$ is either a power series ring $k[[X_0\cdots,X_n]]$ in $n+1$ indeterminates over a field k, or, in the unequal characteristic unramified case, R is a power series ring $S[[X_1,\cdots,X_n]]$ in n indeterminates over the unique unramified complete discrete valuation ring S with residue field k. Following Cohen we call such a ring S simply a v-ring.

In either case the maximal ideal M is generated by $\{X_0,\cdots,X_n\}$ where, in the latter case, X_0 represents p, the characteristic of k.

We choose $\sigma(I)$ to represent the sum $i_0+\cdots+i_n$ where $I=(i_0,\cdots,i_n)$. Let α be in H_r. We write

$$\alpha(A)=A+\sum_{\sigma(I)=r}\alpha_I(A)X^I \tag{2.1}$$

where $\alpha_I(A)$ is some element of R. The quantities $\alpha_I(A)$ are not uniquely determined by α and (2.1). However, they are uniquely determined, mod M, and, one readily verifies, that they have the following properties.

$$\alpha_I(A+B)=\alpha_I(A)+\alpha_I(B),\ \operatorname{mod} M.$$
$$\alpha_I(AB)=A\alpha_I(B)+B\alpha_I(A),\ \operatorname{mod} M.$$

Thus each map α_I induces a derivation d_I on k via the natural map H of R

onto k. We let $\phi(\alpha)$ be the direct product of the $P(\sigma(I)=r)$ derivations d_I ordered lexicographically on $i_0,\cdots,i_n$ where $I=(i_0,\cdots,i_n)$. Thus $P(\sigma(I)=r)$ is the number of I in $\mathfrak{S}$ such that $\sigma(I)=r$. We write G^s for the direct product of the group G to s factors. Thus, if $\mathfrak{D}$ denotes the additive group of derivations on k, a straightforward calculation shows that ϕ is a homomorphism of H_r into $\mathfrak{D}^{P(\sigma(I)=r)}$ with kernel G_{r+1}. We shall show that ϕ is in fact surjective.

As an immediate consequence of the definition of a higher derivation we have

LEMMA 2.1. *If r is a positive integer and $\{D^{(I)}\}$ is a higher derivation on a ring T with the property that $D^{(I)}$ is the zero map for I such that $0<\sigma(I)<r$ then, if $\sigma(J)=r$, $D^{(J)}$ is a derivation.*

In connection with Lemma 2.2 below we make the following observations. Let $d=\{d^{(I)}\}$ be a higher derivation on a commutative ring T with unity. Assume that $d^{(I)}$ is the zero map for I such that $0<\sigma(I)<r$. The mapping $\delta=\delta(d): T\to T[[X]]_n=T[[X_0,\cdots,X_n]]$ given by $\delta(t)=\sum_{I\geqq Q} d^{(I)}(t)X^I$ is an isomorphism of T into $T[[X]]_n$ with the property

$$\delta-\epsilon(T)\subset(X_0,\cdots,X_n)^r \tag{2.2}$$

An isomorphism of T into $T[[X]]_n$ which satisfies (2.2) for $r=1$ we call an embedding of T. The map $d\to\delta(d)$ is a 1-1 correspondence between higher derivations on T and embeddings of T. This correspondence is such that $\delta(d)$ satisfies (2.2) for a given positive integer r if and only if $d^{(I)}$ is the zero map for I such that $0<\sigma(I)<r$.

LEMMA 2.2. *If S is a v-ring with residue field k and $\{d^{(I)}\}$ is a higher derivation on k such that for $0<\sigma(I)<r$, $d^{(I)}$ is the zero map, then there is a higher derivation $\{D^{(I)}\}$ on S which both induces $\{d^{(I)}\}$ and has the property that if $0<\sigma(I)<r$ then $D^{(I)}$ is the zero map.*

Proof. Let H represent the natural map of S onto k and let H' be the extension of H to a homomorphism of $S[[Y]]_n$ onto $k[[X]]_n$ with the property that $H'(Y_i)=X_i$ for $i=0,\cdots,n$. We note that a higher derivation $d=\{d^{(I)}\}$ on k lifts via H to $D=\{D^{(I)}\}$ on S if and only if the corresponding embedding $\delta(d)$ lifts to the embedding $\delta(D)$ via H'. Thus we need to show that an embedding δ on k which satisfies (2.2) for a given positive integer r lifts to an embedding Δ on S which satisfies (2.2) for the same r. This is done by introducing a slight modification in the proof of the embedding theorem of [5, pp. 575-579] and we use the notation of that paper. It is assumed that the embedding θ of k satisfies (2.2) for some r, where, of course, $T = k$. In the construction of the map θ [paragraph 2, p. 575] we always choose a_I (where $H(a_I) = \bar{a}_I$) to be zero if $\bar{a}_I = 0$. Then $\theta - \epsilon(K_1)$ will be in $(Y_1, \cdots, Y_n)^r$. Hence each map θ_m will have the property $\theta_m - \epsilon(R_m) \subset (Y_1, \cdots, Y_n)^r + (p^m, Y_1^{p^m}, \cdots, Y_n^{p^m})$. It follows that θ, where $\theta(a) = \bigcap_{m=1}^{\infty} \theta_m(a + (p^m))$, [p. 579] has the desired property, i.e. $\theta - \epsilon(R) \subset (Y_1, \cdots, Y_n)^r$.

For the remainder of this paper R will be an unramified complete regular local ring of dimension $n+1$ with maximal ideal $M = (X_0, \cdots, X_n)$ where, in the unequal characteristic case, $X_0 = p$ the characteristic of the residue field k.

LEMMA 2.3. *Let r be a positive integer and let $\{A_{I,i}\}^{n}_{\substack{i=0 \\ \sigma(I) \geqq r}}$ be a set of elements in R, arbitrary save for the condition that $A_{I,i}$ is in M if $\sigma(I) = r$ and $i = 0, \cdots, n$. Let $\{d_I\}_{\sigma(I)=r}$ be a set of $P(\sigma(I) = r)$ derivations on k. Then there exists a higher derivation $D = \{D^{(I)}\}$ on R such that,*

(2.3) *$D^{(I)}$ is the zero map if $0 < \sigma(I) < r$,*

(2.4) *$D^{(I)}$ induces d_I for I such that $\sigma(I) = r$,*

(2.5) *$D^{(I)}(X_i) = A_{I,i}$ for $i = 0, \cdots, n$ and I such that $\sigma(I) \geqq r$. If $X_0 = p$ we assume that $A_{I,0} = 0$ for all I.*

Proof. If $R = k[[X_0, \cdots, X_n]]$ we choose a higher derivation $\bar{D} = \{\bar{D}^{(I)}\}$ on k satisfying conditions (2.3) and (2.4). Such a $\bar{D}$ exists by [2, Theorem 1′]if k has characteristic zero and by [3, Theorem 1] otherwise.

An extension D of $\bar{D}$ is then given by the conditions $D = \bar{D}$ on k, $D^{(I)}(X_i) = 0$ if $0 < \sigma(I) < r$ and $D^{(I)}(X_i) = A_{I,i}$ if $\sigma(I) \geqq r$. It follows from [3, Theorem 2] that there is a higher derivation D on R fulfilling the stated conditions.

If $R = S[[X_1, \cdots, X_n]]$, S a v-ring, we appeal to Lemma 2.2 for a higher derivation $\bar{D}$ on S which meets the conditions (2.3) and (2.4). We then define D as above.

LEMMA 2.4. *Let r be a positive integer and let $D = \{D^{(I)}\}$ be a higher derivation on R having the following properties.*

(2.6) *$D^{(I)}$ is the zero map for I such that $0 < \sigma(I) < r$.*

(2.7) *If $\sigma(I) = r$ then $D^{(I)}(X_i) \in M$ for $i = 0, \cdots, n$.*

Then the map α_D given by

$$\alpha_D(A) = \sum_{I \geqq Q} D^{(I)}(A) X^I$$

is in H_r.

Proof. By definition of a higher derivation α_D is a homomorphism and is in H_r if it is an automorphism. Since $\alpha_D(A) = A$, mod M, it follows that α_D is surjective. A calculation shows that if A is in M^s so is $D^{(J)}(A)$ for all $J \geqq Q$. Moreover, if A is in M^s and not in M^{s+1} then $\alpha_D(A) - A$ is in M^{s+1}. Thus $\alpha_D(A) = 0$ implies $A = 0$.

THEOREM 2.1. *If R is an unramified complete regular local ring of dimension $n+1$ then, for $r = 1, 2, \cdots$ H_r/G_{r+1} is isomorphic to $\mathfrak{D}^{P(\sigma(I)=r)}$, the direct product of the additive group of derivations $\mathfrak{D}$ on the residue field of R to $P(\sigma(I) = r)$ factors where $P(\sigma(I) = r)$ is the number of ordered partitions of the integer r into non-negative summands.*

COROLLARY 2.1. *The equality $H_r = G_r$ holds for $r = 1, 2, \cdots$ if and only if k is algebraic over the field of rationals or k is perfect, in the case in which k has characteristic $p \neq 0$.*

Proof. If k has characteristic zero $\mathfrak{D}$ is trivial if and only if k is algebraic over the field of rationals and in the remaining case $\mathfrak{D}$ is trivial if and only if k is perfect.

THEOREM 2.2. *G_1/H_1 is isomorphic to the full linear group $L_{n+1}(k)$ in the equal characteristic case and is isomorphic to the subgroup of $L_{n+1}(k)$ with first row $1, 0, \cdots, 0$ in the unequal characteristic case.*

Proof. We first consider the equal characteristic case. For α in G_1, there are elements a_{ij} in k such that $\alpha(X_i) = \sum_{j=0}^{n} a_{ij}X_j$, mod M^2. Necessarily $\det(a_{ij}) \neq 0$. The map which takes α onto the matrix $A = (a_{ij})$ is easily seen to be a homomorphism of G_1 onto $L_{n+1}(k)$ with kernel H_1.

In the unequal characteristic case we again have $\alpha(X_i) = \sum_{j=0}^{n} a_{ij}X_j$, mod M^2, for $i = 0, \cdots, n$. However, in this case, the a_{ij} are in S and $\alpha(X_0) = X_0$. Writing $\bar{a}_{ij}$ for the image of a_{ij} by the natural map of S onto k the correspondence $\alpha \rightarrow (\bar{a}_{ij})$ maps G_1 homomorphically onto the subgroup of $L_{n+1}(k)$ with first row $1, 0, \cdots, 0$ and kernel H_1.

THEOREM 2.3. *If R has dimension $n+1$ and $r > 1$, then G_r/H_r is isomorphic to $(k^+)^{(n+1)(P(\sigma(I)=r))}$, the direct product of the additive group k^+ of the residue field with itself to $(n+1)(P(\sigma(I) = r))$ factors.*

Proof. We consider the equal characteristic case first. Let α be in G_r. Then, for $i = 0, \cdots, n$, $\alpha(X_i) = X_i + \sum_{\sigma(I)=r} a_{I,i}X^I$, mod M^{r+1}, where $a_{I,i}$ is in k. The map $\alpha \rightarrow \{a_{I,i}\}^{n}_{i=0,\sigma(I)=r}$ is a homomorphism of G_r onto $(k^+)^{(n+1)P(\sigma(I)=r)}$ with kernel H_r.

In the unequal characteristic case we argue similarly using the fact that for $i=1,\cdots,n$

$$\alpha(X_i)=X_i+\sum_{j=0}^{r}p^j(\sum_{\substack{I=(0,i_1,\cdots,i_n)\\ \sigma(I)=r-j}}a_{I,i}X^I),\ \operatorname{mod}M^{r+1},$$

where the elements $a_{I,i}$ are in S. In this case the map $\alpha\rightarrow\{\bar{a}_{I,i}\}$ yields the desired result.

III. Representations of inertial automorphisms by derivations and higher derivations. Let $D=\{D^{(J)}\}$ be a higher derivation on R. As we have noted before, the map α_D given by

$$\alpha_D(A)=\sum_{J\geqq Q}D^{(I)}(A)X^I \tag{3.1}$$

is a homomorphism. Let I_i represent the $n+1$ tuple with 1 in the i-th position and zeros elsewhere. Let a_{ij} be the image of $D^{(I_i)}(X_j)+\delta_{ij}$ (δ_{ij} representing "Kronecker delta") under the natural map of R onto k. Now α_D will be an automorphism and hence an element of G_1 if and only if $\det(a_{ij})$, which we call the *determinant of D,* is different from zero. We will show in this section that every α in G_1 has the form (3.1) from which fact it will follow, if k has characteristic zero or R is a v-ring, that every α in G_1 has a Taylor series like representation in terms of a sequence of derivations $\{D_I\}_{I>Q}$ on R.

Let α be in G_1. We choose elements $\alpha'_j(X_i)$ in R so that

$$\alpha(X_i)=X_i+\sum_{j=0}^{n}\alpha'_j(X_i)X_j$$

for $i=0,\cdots,n$. If $X_0=p$, we let $\alpha'_j(X_0)=0$ for all j.

By the proof of Lemma 2.3 there exists a higher derivation $D_0=\{D_0{}^{(I)}\}$ on R such that

(i) $D_0{}^{(I_j)}(X_i)=\alpha'_j(X_i)$ for $i,j=0,1,\cdots,n$.

and

(ii) If $\sigma(I)>1$ then $D_0{}^{(I)}(X_i)=0$ for $i=0,\cdots,n$.

Let α_{D_0} be defined by D_0 as in (3.1). Then, by choice of D_0 it follows that α_{D_0} is in G_1, $\alpha_1=\alpha\alpha_{D_0}{}^{-1}$ is in H_1 and $\alpha_1(X_i)=X_i$ for $i=0,\cdots,n$.

Assume that r higher derivations $D_i=\{D_i{}^{(I)}\}$ have been obtained so that $\alpha_{i+1}=\alpha_i\alpha_{D_i}{}^{-1}$ is in H_{i+1} for $i=0,\cdots,r-1$; and for $i\geqq 1$, $n\geqq j\geqq 0$, $D_i{}^{(J)}(X_j)=0$, for all $J>Q$. For each I such that $\sigma(I)=r$ we choose maps $\alpha'_I\colon R\rightarrow R$ so that $\alpha_r(A)=A+\sum_{\sigma(I)=r}\alpha'_I(A)X^I$ for all A in R. Again, by Lemma 2.3, there exists a higher derivation D_r on R having the following properties:

(a) If $0 < \sigma(I) < r$ then $D_r^{(I)}$ is the zero map.

(b) If $\sigma(I) = r$ then $D_r^{(I)}(A) = \alpha'_I(A)$, $\mod M$.

(c) $D_r^{(I)}(X_i) = 0$ for $I > Q$ and $i = 0, \cdots, n$.

The automorphism α_{D_r} defined by (3.1) using such a higher derivation D_r will have the properties; $\alpha_{r+1} = \alpha_r \alpha_{D_r}^{-1}$ is in H_{r+1} and $\alpha_{r+1}(X_i) = X_i$ for $i = 0, \cdots, n$.

In this manner we obtain a sequence of automorphisms α_{D_i} for $i = 0, \cdots$ with the property that

$$\alpha = \lim_{n \to \infty} \alpha_{D_n} \alpha_{D_{n-1}} \cdots \alpha_{D_0}. \tag{3.2}$$

where the limit map (3.2) takes A into $\lim\limits_{n \to \infty} A_n$ and $A_n = \alpha_{D_n} \cdots \alpha_{D_0}(A)$. Since α_{D_i} is in H_i $(i > 0)$ it is clear that the limit (3.2) exists.

As a consequence of (3.2) we write

$$\alpha(A) = \sum_{I \geqq Q} D_\alpha^{(I)}(A) X^I$$

where, for I such that $\sigma(I) = r > 0$,

$$D_\alpha^{(I)} = \sum_{\substack{I_0, \cdots, I_r \\ I_0 + \cdots + I_r = I}} D_r^{(I_r)} D_{r-1}^{(I_{r-1})} \cdots D_0^{(I_0)} \tag{3.3}$$

It is readily shown that $D_\alpha = \{D_\alpha^{(I)}\}$ as defined by (3.3) is a higher derivation. Thus we have a converse to Lemma 2.4 and putting the two together we obtain

THEOREM 3.1. *Let R be an unramified complete regular local ring of dimension $n+1$ having maximal ideal $(X_0, \cdots, X_n)$. Let $D = \{D^{(I)}\}$ be a higher derivation on R with non-zero determinant. Then, the mapping $\alpha_D \colon A \to \sum_{I \geqq Q} D^{(I)}(A) X^I$ is an inertial automorphism and, conversely, given an inertial automorphism α on R, there exists a higher derivation D on R such that $\alpha = \alpha_D$. Necessarily D will have non-zero determinant.*

An explicit description of inertial automorphisms in terms of derivations is now available in some cases. In [2, Proof of Theorem 3′] it is shown that if C is a commutative ring containing the field of rational numbers and $\{D_I\}_{I > Q}$ is a sequence of derivations on C then the set of maps $\{D^{(I)}\}$ where $D^{(I)} = \sum_{(r,I)} \dfrac{[D]_{(r,I)}}{r!}$ is a higher derivation. (Here (r, I) is a partition of I into r non-zero summands $I_1, \cdots, I_r$ and $[D]_{(r,I)}$ represents the sum of all the formally distinct products of the r derivations $D_{I_1}, \cdots, D_{I_r}$). Moreover, every higher derivation on D is obtained in this manner.

Thus, if k has characteristic zero R is such a ring C and we have

THEOREM 3.2. *Let k be a field of characteristic zero and let $\{D_I\}_{I>Q}$ be a sequence of derivations on the power series ring $R = k[[X_0, \cdots, X_n]]$. If the determinant of the matrix $(A_{i,j})$ where $A_{i,j} = D_{I_i}(X_j) + \delta_{ij}$ is not in the maximal ideal of R then the map α given by*

$$\alpha(A) = A + (\sum_{I>Q} (\sum_{(r,I)} \frac{[D]_{(r,I)}(A)}{(r!)}) X^I$$

is an inertial automorphism of R. Moreover, every inertial automorphism of R has such a representation.

If R is a v-ring and D_s is a derivation on R we let $\alpha_{D_s,i}$ represent the automorphism defined by

$$\alpha_{D_s,i}(A) = A + \sum_{n=1}^{\infty} \frac{D_s^n(A)}{n!} p^{ni} \tag{3.4}$$

where D_s^n is the n-th iterate of the derivation D_s. If either $p \neq 2$ or $i > 1$, then (3.4) is an automorphism in $G_i = H_i$. Thus, by the argument preceding Theorem 3.1, we conclude that if α is in G_i then

$$\alpha = \lim_{n \to \infty} \alpha_{D_n, i+n} \cdots \alpha_{D_1, i+1} \alpha_{D_0, i} \tag{3.5}$$

for appropriate choice of derivations $D_0, D_1, \cdots$.

THEOREM 3.3. *Let R be a v-ring. If α is in G_i and, $i > 1$, or p, the characteristic of the residue field, is different from 2, then there is a sequence of derivations $\{D_i\}_{i=0}^{\infty}$ on R such that for all A in R*

$$\alpha(A) = \sum_{s=0}^{\infty} \sum_{t_i=0}^{\infty} \frac{D_s^{t_s} D_{s-1}^{t_{s-1}} \cdots D_0^{t_0}(A)}{t_s! \cdots t_0!} p^{t_0(i) + t_1(i+1) + \cdots + t_s(i+s)}. \tag{3.6}$$

Conversely, given a sequence $\{D_i\}_{i=0}^{\infty}$ of derivations on R the mapping α defined by (3.6) is in G_i.

IV. Inertial automorphisms and compositions on higher derivations. The composition $C \circ D$ of higher derivations C and D given by (3.3) in which $(C \circ D)^{(I)} = \sum_{Q \leq J \leq I} C^{(J)} D^{(I-J)}$ makes a group out of the set $\mathfrak{H}$ of higher derivations on R. This can be shown directly or by using the one-to-one correspondence between elements of $\mathfrak{H}$ and automorphisms β on $R[[X]]_n$ such that $\beta - \epsilon(R) \subset (X_0, \cdots, X_n)$ and $\beta(X_i) = X_i$ for $i = 0, \cdots, n$. However, the condition $\alpha_{C \circ D} = \alpha_C \alpha_D$ is fulfilled only if $\alpha_C(X_i) = X_i$ for $i = 0, \cdots, n$.

To obtain a composition which is preserved under the map $D \rightarrow \alpha_D$ we note that $\alpha_C\alpha_D(A) = \sum_{I \geqq Q} \alpha_C(X^I) \sum_{J \geqq Q} C^{(J)}D^{(I)}(A)X^J$. If we write $\alpha(X)^I$ for $\alpha(X_0)^{i_0}\alpha(X_1)^{i_1} \cdots \alpha(X_n)^{i_n}$ where $I = i_0, \cdots, i_n$ and Let A_J represent the coefficient of X^J in the power series A, we have,

$$\alpha_C\alpha_D(A) = \sum_{K \geqq Q} \Bigl(\sum_{\substack{K \geqq J \geqq Q \\ I \geqq Q}} [\alpha_C(X)^I]_{K-J} C^{(J)}D^{(I)}(A) \Bigr) X^K. \tag{4.1}$$

Relation (4.1) suggests the following.

THEOREM 4.1. *Let $P_0, \cdots, P_n$ be $n+1$ elements in M. Let $C = \{C^{(I)}\}$ and $D = \{D^{(I)}\}$ be higher derivations on R. Then $E = \{E^{(I)}\}$ where*

$$E^{(K)} = \sum_{I > Q} \sum_{K \geqq J \geqq Q} [P^I]_{K-J} C^{(J)}D^{(I)} \tag{4.2}$$

is a higher derivation.

Proof. The map $E^{(K)}$ is clearly additive. We need to show that

$$E^{(K)}(AB) = \sum_{K \geqq J \geqq Q} E^{(K-J)}(A)E^{(J)}(B)$$

Now

$$E^{(K)}(AB)$$
$$= \sum_{I > Q} \sum_{K \geqq J \geqq Q} [P^I]_{K-J} \sum_{\substack{J \geqq J_1 \geqq Q \\ I \geqq I_1 \geqq Q}} C^{(J_1)}D^{(I_1)}(A)C^{(J-J_1)}D^{(I-I_1)}(B), \tag{4.3}$$

and

$$\sum_{K \geqq K_1 \geqq Q} E^{(K)}(A)E^{(K-K_1)}(B) \tag{4.4}$$
$$= \Bigl[\sum_{I_1 \geqq Q} \sum_{K_1 \geqq J_1 \geqq Q} [P^{I_1}]_{K_1-J_1} C^{J_1}D^{I_1}(A) \Bigr] \Bigl[\sum_{I \geqq Q} \sum_{K-K_1 \geqq J \geqq Q} [P^I]_{K-K_1-J} C^{(J)}D^{(I)}(B) \Bigr]$$
$$= \sum_{I \geqq Q} \sum_{K \geqq J \geqq Q} \sum_{\substack{J \geqq J_1 \geqq Q \\ I \geqq I_1 \geqq Q}} \Bigl[\sum_{\substack{K-J \geqq \\ K_1-J_1 \geqq \\ Q}} [P^{I_1}]_{K_1-J_1} [P^{I-I_1}]_{K-J-(K_1-J_1)} \Bigr] C^{J_1}D^{I_1}(A)C^{J-J_1}D^{I-I_1}(B).$$

But, by definition of the product of power series

$$\sum_{K-J \geqq K_1-J_1 \geqq Q} [P^{I_1}]_{K_1-J_1} [P^{I-I_1}]_{K-J-(K_1-J_1)} = [P^I]_{K-J},$$

from which fact we see that (4.4) is the same as (4.3). We let $C * D$ be the higher derivation E obtained via (4.2) in that case in which $P_i = \alpha_C(X_i)$. In view of (4.1) we have $\alpha_{C*D} = \alpha_C\alpha_D$. The writer has not determined what sort of algebraic structure the system $\{\mathfrak{H}, *\}$ has, except that it is non associative. However, in the equal characteristic case we have the following.

THEOREM 4.2. *Let k be a field and let $R = k[[X]]_n$. Let $\mathfrak{H}_k$ be the set of higher derivations $D = \{D^{(I)}\}$ on R such that*

(i) *$D^{(I)}(k) \subset k$, for $I > Q$,*

(ii) *$D^{(I)}(X_i) \in k$, for $I > Q$, and, $i = 0, \cdots, n$,*

(iii) *D has non-zero determinant.*

The map $D \to \alpha_D$ is an isomorphism between $\{\mathfrak{H}_k, *\}$ and G_1.

Proof. We show first that $D \to \alpha_D$ yields a biunique surjective correspondence between $\mathfrak{H}_k$ and G_i. The uniqueness of the coefficients in a power series imply that if $\alpha_D = \alpha_E$ for D and E in $\mathfrak{H}_k$ then $D = E$.

Let α be an inertial automorphism on R. The restriction α_k of α to k is an embedding of k in $k[[X]]_n$. If $\alpha_k(a) = \sum_{I \geqq Q} a_I X^I$ where $a_I \in k$, then the higher derivation $D_k = \{D_k^{(I)}\}$ given by $D_k^{(I)}(a) = a_I$ is clearly such that $\alpha_k(a) = \sum_{I \geqq Q} D_k^{(I)}(a) X^I$. If $\alpha(X_i) = X_i + \sum_{I > Q} a_{i,I} X^I$ we extend D_k to $D = \{D^{(I)}\}$ on R, as in the proof of Lemma 2.3 by the condition $D^{(I)}(X_i) = a_{i,I}$ for all $I > Q$ and $i = 0, \cdots, n$. We observe, by appeal to the definition of $*$, that if C and D are in $\mathfrak{H}_k$ then $C * D$ is in $\mathfrak{H}_k$. The relation $\alpha_{C*D} = \alpha_C \alpha_D$ completes the proof.

A similar result does not seem to be available in the unequal characteristic case since R does not then have a sufficiently well behaved set of representatives of the residue field.

FLORIDA STATE UNIVERSITY.

REFERENCES.

[1] I. S. Cohen, "Structure of complete local rings," *Transactions of the American Mathematical Society*, vol. 59 (1946), pp. 54-106.

[2] N. Heerema, "Derivations and embeddings of a field in its power series ring," *Proceedings of the American Mathematical Society*, vol. 11 (1960), pp. 188-194.

[3] ———, "Derivations and embeddings of a field in its power series ring, II," *Michigan Mathematical Journal*, vol. 8 (1961), pp. 129-134.

[4] ———, "Derivations on p-adic fields," *Transactions of the American Mathematical Society*, vol. 102 (1962), pp. 346-351.

[5] ———, "Embeddings of a p-adic field and its residue field in their power series rings," *Proceedings of the American Mathematical Society*, vol. 14 (1963), pp. 37-43.

[6] S. MacLane, "Subfields and automorphism groups of p-adic fields," *Annals of Mathematics*, vol. 40 (1939), pp. 423-442.

G-INDUCED FLOWS.

By L. Auslander and L. W. Green.[1]

1. Introduction. Let G be a connected, simply-connected Lie group and C a closed subgroup such that G/C has a finite invariant measure. If $\{g_t\}$ is a one-parameter subgroup of G, it induces a one-parameter group of measure-preserving diffeomorphisms of G/C, called a G-induced flow. The questions we consider here concern the dynamics of such flows; namely, when is a G-induced flow ergodic, and when does there exist an ergodic G-induced flow?

The case when G is commutative can now be considered "classical" in ergodic theory (see, for a brief discussion and generalizations in a direction different from ours, Mackey [9a]). For special kinds of solvable G, including all nilpotent groups, the problem was solved in [1]. In this paper we reduce the general problem to cases already treated.

The principal concept upon which the reduction hinges is that of a type (R) Lie algebra. A real Lie algebra is called type (R) if no element in the adjoint representation has a characteristic root with non-zero real part. A Lie group G, as well as its homogeneous spaces, will be called type (R) if the Lie algebra of G is type (R). Compact groups, nilpotent groups, and certain solvable groups (like the rigid motions of the plane, hence the name) are type (R).

The first reduction (Theorem 3.5) is effected by proving that there exists a closed subgroup K containing C such that G/K is a type (R) homogeneous space and on which the associated flow is ergodic if and only if the original flow was. In the interesting case when the semi-simple part of G has no compact simple factors, and the flow is general enough, G/K turns out to be a compact solvmanifold. The second reduction (Theorem 4.5) shows that flows on compact solvmanifolds of type (R) are equivalent to flows on associated nilmanifolds, and can thus be assumed analyzed in [1]. A consequence is that in general, there always exist G-induced ergodic flows.

It is undoubtedly true (we do not prove it here) that the flow on the

---

Received May 22, 1964.

Revised September 8, 1965.

[1] Research supported in part by National Science Foundation Grants G 24336, GP 3750.

type (R) solvmanifold to which we reduce the problem is a maximal distal subflow of the original, in the terminology of Furstenberg [7]. The general case represents, then, a class of ergodic, non-isometric, non-minimal extensions of distal minimal flows.

In Appendix I, the first author proves a converse of the result on type (R) solvmanifolds: namely, a compact solvmanifold with nilpotent fundamental group is of type (R). In Appendix II, the second author gives an elementary proof of part of "Mautner's Lemma" (so-called in [1]), thus making most of the reduction steps independent of any extensive theory of group representations.

We are indebted to the referee for many clarifying suggestions, and to C. C. Moore for pointing out that our original version, which treated only solvable G, could be extended to the general case very easily. The results for the solvable case were announced in [2]. Special cases had been studied in [3] and [8]. In a forthcoming paper, "Ergodicity of flows on homogeneous spaces," C. C. Moore gives a much more exhaustive treatment of the semisimple case.

Throughout the paper the isotropy group C will be assumed to contain no connected subgroups normal in G.

2. Let $\mathfrak{G}$ be a real Lie algebra, $X \in \mathfrak{G}$. We decompose $\mathfrak{G}$ as a vector space into primary components relative to $\operatorname{ad} X$:

$$\mathfrak{G} = \mathfrak{G}_0 + \mathfrak{G}_1 + \cdots + \mathfrak{G}_m.$$

Here each $\mathfrak{G}_i$ is invariant under $\operatorname{ad} X$, and the minimal polynomial of the restriction of $\operatorname{ad} X$ to $\mathfrak{G}_i$ is an irreducible polynomial $\mu_i(\lambda)$. $\mathfrak{G}_i$ is the set of elements Z such that $\mu_i(\operatorname{ad} X)^r Z = 0$ for some r. We reserve the index 0 for the null component of $\operatorname{ad} X$, namely, $\mu_0(\lambda) = \lambda$. Since we are using the field of real numbers, the possibilities for μ_i are limited to linear and quadratic polynomials. In analogy with the terms used in the qualitative theory of differential equations, we call $\mathfrak{G}_i$ an *unstable component* and $\mu_i(\lambda)$ an *unstable polynomial* if the roots of $\mu_i(\lambda) = 0$ have non-vanishing real parts. The smallest subalgebra containing all the unstable components will be called the subalgebra unstable relative to X. (This subalgebra may contain parts of "stable" primary components, or even be all of $\mathfrak{G}$.)

For convenience in the statement of the next lemma, we let $\mathfrak{G}_i$ designate the Fitting null component of $\mu_i(\operatorname{ad} X)$, whether $\mu_i(\lambda)$ is a primary function or not. In the latter case, $\mathfrak{G}_i$ is, of course, the zero subspace.

LEMMA 2.1. *Let the roots of* $\mu_i(\lambda)$, $\mu_j(\lambda) = 0$ *be* ω, $\bar{\omega}$, γ, $\bar{\gamma}$ *respectively. Set*

$$\mu_k(\lambda) = (\lambda - (\omega + \gamma))(\lambda - (\bar{\omega} + \bar{\gamma})),$$

$$\mu_s(\lambda) = (\lambda - (\omega + \bar{\gamma}))(\lambda - \bar{\omega} + \gamma)), \text{ and } \mu_t(\lambda) = \lambda - \omega + \gamma).$$

Then

i) *if* ω *and* γ *are both real,*

$$[\mathfrak{G}_i, \mathfrak{G}_j] \subset \mathfrak{G}_t;$$

ii) *if* $\omega = -\gamma$, *and is not real,*

$$[\mathfrak{G}_i, \mathfrak{G}_j] \subset \mathfrak{G}_s + \mathfrak{G}_t;$$

iii) *if* $\omega \neq -\gamma, -\bar{\gamma}$, *and one of these three numbers is not real;*

$$[\mathfrak{G}_i, \mathfrak{G}_j] \subset \mathfrak{G}_k + \mathfrak{G}_s.$$

Proof. The proof is immediate from considerations of the complexification $\mathfrak{L}$ of $\mathfrak{G}$. Namely, $\mathfrak{G}_i$, for example, may be identified with the real vectors in $\mathfrak{L}_\omega + \mathfrak{L}_{\bar{\omega}}$, where $\mathfrak{L}_\omega$ is the component corresponding to the complex primary function $\lambda - \omega$. The standard formula

$$[\mathfrak{L}_\alpha, \mathfrak{L}_\beta] \subset \mathfrak{L}_{\alpha+\beta}$$

may then be applied and the conclusions of the lemma read off.

COROLLARY 2.2. *The subalgebra unstable relative to* X *is an ideal.*

Proof. Let $\mathfrak{U}$ denote the unstable subalgebra. It is sufficient to show that, if $Y \in \mathfrak{G}_j$ and $Z \in \mathfrak{G}_i$, with $\mathfrak{G}_i$ an unstable component, then $[Y, Z] \in \mathfrak{U}$. Let ω be a root corresponding to $\mathfrak{G}_i$, and consequently having non-zero real part. If γ is a root corresponding to $\mathfrak{G}_j$, the lemma tells us that $[Y, Z]$ will be in a sum of unstable components unless possibly the real part of γ is the negative of the real part of ω. But then $\mathfrak{G}_j$ was already unstable, and $[Y, Z] \in \mathfrak{U}$ by definition.

We recall the definitions of two real solvable Lie algebras which turn up repeatedly in contexts like this. $\mathfrak{S}_1$ is the two-dimensional algebra with basis U, V and relation $[U, V] = V$. $\mathfrak{S}_2$ has basis U, V, W and relations $[U, V] = V - \sigma W$, $[U, W] = \sigma V + W$, $[V, W] = 0$, where σ is a non-zero real number. Relative to U each of these algebras contains a single unstable component, and the primary functions are, respectively, $\lambda - 1$ and $\lambda^2 - 2\lambda + \sigma^2 + 1$. Any non-zero multiple of U in either of these algebras will be called regular.

LEMMA 2.3. *Let* $\mathfrak{A}$ *be the sum of all unstable components relative to* X *belonging to primary functions with roots having strictly positive real parts. Then* $\mathfrak{A}$ *is a nilpotent subalgebra and contains a sequence* $\{\mathfrak{A}_i\}$ *of subspaces such that*

i) $\mathfrak{A} = \mathfrak{A}_0 \supset \mathfrak{A}_1 \supset \cdots \supset \mathfrak{A}_r$, *and each* $\mathfrak{A}_i$ *is invariant under* $\operatorname{ad} X$;

ii) $\mathfrak{A}_{i+1}$ *is an ideal in* $\{X\} + \mathfrak{A}_i$;

iii) $\{X\} + \mathfrak{A}_i/\mathfrak{A}_{i+1}$ *is isomorphic to either* $\mathfrak{S}_1$ *or* $\mathfrak{S}_2$, *with the image of* X *regular.*

Outline of proof. This lemma is essentially Theorem 2.2 of [1], page 75, applied to the almost type (E) solvable algebra $\{X\} + \mathfrak{A}$. The fact that X was regular in the entire algebra was not used there. The inductive proof begins by splitting an unstable component corresponding to a root with maximal real part to form the bottom of the sequence of subalgebras (see Lemma 2.1 of [1], page 75). Repeated use is made of Lemma 2.1, above, to prove ii). The decomposition is not unique, since, modulo relations between the roots, the ordering of the unstable components is arbitrary.

The Mautner Lemma (see Appendix II) states that in any unitary representation of a group with algebra $\mathfrak{S}_i$, eigenvectors of the image of $\exp(tU)$ are fixed by the images of $\exp V$, $\exp W$. Combining this with Lemma 2.3, we may prove

THEOREM 2.4. *Let* $g \to U(g)$ *be a continuous unitary representation of a connected Lie group whose algebra is* $\mathfrak{G}$. *If* $U(\exp tX)\psi = e^{i\lambda t}\psi$ *for some* λ *and all* t, *then* $U(\exp Y)\psi = \psi$ *for every* Y *in the ideal of* $\mathfrak{G}$ *unstable relative to* X.

The proof is the same as that of Theorem 3.2, [1], page 76. After proving the theorem for all $Y \in \mathfrak{A}$ by climbing up the ladder of subalgebras, one does the analogous thing for $\mathfrak{B}$, the subalgebra corresponding to roots with strictly negative real parts (equivalently, the $\mathfrak{A}$ for $-X$). $\mathfrak{A}$ and $\mathfrak{B}$ generate the ideal of $\mathfrak{G}$ unstable relative to X.

3. To apply the results of the previous section to ergodic theory, it is desirable to have the unstable ideal as large as possible. (This is not a necessary condition for ergodicity, however; witness the horocycle flow.) To this end we instroduce some terminology.

Definitions. An element X of a real Lie algebra is *really unstable* if the number of characteristic values with non-zero real part of $\operatorname{ad} X$ is maximal. X is called *really regular* if it is regular and really unstable. A real Lie algebra is said to be of *type* (R) if no element has a non-trivial unstable component.

In this last definition, it is sufficient to require that no really regular element has a non-trivial unstable component; for the really regular elements form an open dense set in the algebra and the presence of a characteristic value off the pure imaginary axis persists under small perturbations. Notice that the image of a really regular element under a homomorphism is really regular (proof as in Chevalley [6a], page 217).

LEMMA 3.1. *A semi-simple type (R) Lie algebra is compact.*

Proof. Let X be really regular ($=$ regular) in the semi-simple algebra $\mathfrak{G}$, and let $\mathfrak{H}$ be the Cartan subalgebra of $\mathfrak{L}$, the complexification of $\mathfrak{G}$, which contains X. Every root of $\mathfrak{H}$ must be of the form $i\lambda$, where λ is a real (relative to $\mathfrak{H} \cap \mathfrak{G}$) linear form, otherwise X would have a non-trivial stable component. Thus the Killing form is negative definite, and $\mathfrak{G}$ is compact.

LEMMA 3.2. *The homomorphic image of a type (R) algebra is type (R).*

Proof. If σ denotes both the homomorphism and the mapping induced on the adjoint algebra, then $\operatorname{ad}(X^\sigma) = (\operatorname{ad} X)^\sigma$, and the minimal polynomial of $(\operatorname{ad} X)^\sigma$ is a factor of that of $\operatorname{ad} X$. Thus no unstable polynomial can arise if there were none to begin with. (This also begins a proof of the above remark about the images of really regular elements, modulo regularity.)

THEOREM 3.3. *Let $\mathfrak{R}$ be the radical of the real Lie algebra $\mathfrak{G}$ and suppose $\mathfrak{G}/\mathfrak{R}$ has no compact ideals. If $\mathfrak{U}$ is the unstable ideal relative to a really regular element, $\mathfrak{G}/\mathfrak{U}$ is a type (R) image of $\mathfrak{R}$, hence solvable.*

Proof. Clearly $\mathfrak{G}/\mathfrak{U}$ is of type (R). Map it onto $\mathfrak{G}/(\mathfrak{U}+\mathfrak{R})$ by the natural homomorphism to obtain another type (R) algebra, by Lemma 3.2. But the latter, being also an image of $\mathfrak{G}/\mathfrak{R}$, is semi-simple with no compact ideals, hence trivial, by Lemma 3.1. Therefore $\mathfrak{U}+\mathfrak{R}=\mathfrak{G}$, and

$$\mathfrak{G}/\mathfrak{U} = (\mathfrak{U}+\mathfrak{R})/\mathfrak{U} \cong \mathfrak{R}/(\mathfrak{U}\cap\mathfrak{R}).$$

This completes the proof of the theorem.

It is possible that $\mathfrak{R}/(\mathfrak{U}+\mathfrak{R})$ will be smaller than the (unique within isomorphism) largest type (R) image of $\mathfrak{R}$. This of course only increases the chance of inducing an ergodic flow.

The following lemma is contained in the proof of Theorem 7.2 in Mackey [9], as was pointed out to us by C. C. Moore.

LEMMA 3.4. *Let f be a bounded Borel function on the locally compact group G, and set $(D_x f)(g) = f(gx)$. Then the set of g such that $D_g f = f$ almost everywhere is a closed subgroup of G.*

Proof. Introduce right invariant Haar measure μ and consider f as a member of $L_\infty(G,\mu)$. The point is that, although $g \to D_g$ is a strongly continuous representation of G in $L_p(G)$ for $1 \leqq p < \infty$, this is not the case for $p = \infty$. However, it is continuous in the L_1-topology of L_∞. Namely, for $h \in L_1$ and $f \in L_\infty$,

$$\int h D_g f = \int (D_g^{-1} h) f,$$

and represents a continuous scalar valued function of g. If $C(h)$ is the closed set of g's for which this number equals $\int hf$, the subgroup in question is the intersection of all $C(h)$ as h runs over L_1.

The situation to which we wish to apply these results is the following: C is a closed subgroup of the connected Lie group G such that G/C has a finite G-invariant measure, μ. The map $g \to \Gamma_g$, where $(\Gamma_g \psi)(hC) = \psi(g^{-1}hC)$ for $\psi \in L_2(G/C,\mu)$, is a continuous unitary representation of G (the representation induced by the identity representation of C). Let X be in the Lie algebra of G and U be the analytic normal subgroup corresponding to the unstable ideal relative to X. The restriction to the one-parameter subgroup $\{\exp tX \mid -\infty < t < \infty\}$ of the action of G on G/C is called a *flow.* If $\Gamma(\exp tX)\psi = e^{t\lambda t}\psi$ for some λ and all t, ψ is called an eigenvector of the flow. When the only eigenvectors (eigenvectors for $\lambda = 0$) are the constant functions, the flow is called weakly mixing (resp., ergodic). Set $K = \overline{CU}$ and denote the natural map of G/C on G/K by ϕ.

THEOREM 3.5. *Every eigenvector of the flow on G/C is of the form $\Psi \cdot \phi$, where Ψ is an eigenvector of the corresponding flow on G/K. In particular, the flow on G/C is ergodic (mixing) if and only if the flow on G/K is.*

Proof. Let ψ be an eigenvector of the flow. By Theorem 2.4, $\Gamma_u \psi = \psi$ for every $u \in U$. Without loss of generality we may assume ψ is essentially bounded (the truncation of an eigenvector is an eigenvector) and lift it to an element in $L_\infty(G)$. A representation f for this element may be chosen such that, for all g outside of a fixed null set of G, $f(ug) = f(g) = f(gc)$ for every $u \in U$, $c \in C$. If $u' \in U$, $f(gu') = f(ug) = f(g)$, where $u = gu'g^{-1}$. Thus Lemma 3.4 applies, and the closed subgroup of h's for which $D_h f = f$ certainly contains K. Thus f is the lift of an eigenvector of the flow on G/K, and so is ψ, as was to be proved.

If X is really regular, Theorem 3.3 implies that G/K is a solvmanifold, being a homogeneous space of G/U. We call G/K a type (R) solvmanifold, since G/U is type (R). Now μ is a finite G-invariant measure on G/C, so its image under ϕ is G/U-invariant on G/K. Thus a theorem of Mostow ([13], Theorem 6.2) applies, and we conclude that G/K is compact.

COROLLARY 3.6. *Let R be the radical of G and suppose G/R has no compact simple factor. Then the flow on G/C defined by a really regular element is ergodic if and only if the corresponding flow on the compact type (R) solvmanifold G/K is.*

If G/K is non-trivial, the possibility of mixing is excluded by the analysis of the remainder of this paper, where the ergodic and topological properties of the flow on G/K are shown to be equivalent to those of an associated nilflow. A nilflow is never mixing. However, when G is semi-simple, we may state

COROLLARY 3.7. *Let G/C be a homogeneous space with finite invariant measure of the semi-simple Lie group G. If G has no compact simple factors, any flow on G/C defined by a really unstable element is mixing.*

For the proof, one merely notices that the unstable ideal relative to a really unstable element must have a non-trivial intersection with each non-compact direct summand.

4. Let G be a connected, simply connected solvable Lie group and let the maximal connected, nilpotent normal subgroup of G be called the nil-radical. It will often be convenient to identify a connected, simply connected nil-potent Lie group with its Lie algebra by means of the exponential mapping. This identification will be made implicitly whenever it is convenient. We will also use solvable Lie group to denote a connected and simply connected solvable Lie group. Our object in this section is to discuss a useful modification of the concept of splitting algebra introduced by Malcev in [11].

Let G be a solvable Lie group with nil-radical N. Then if R^r denotes r copies of the reals, we have the exact diagram

$$1 \to N \to G \to R^r \to 1$$

Let $\mathrm{ad}(G)$ denote the adjoint representation of G. Then $\mathrm{ad}(G)$ is a matrix group. We may consider the intersection of all the real algebraic matrix groups containing $\mathrm{ad}(G)$. This group is the algebraic hull of $\mathrm{ad}(G)$ and will be denoted by $G^\#$. Then by Borel [5] we see that $G^\#$ is solvable and connected, since $\mathrm{ad}(G)$ is connected. We also find in Borel [5] a proof of the crucial fact that

$$G^{\#} = T^{\#} \cdot N^{\#}$$

where $T^{\#}$ is an abelian semi-simple matrix group (semi-simple = completely reducible) and $N^{\#}$ is the maximal unipotent (all eigen-values are one) subgroup of $G^{\#}$. Clearly $N^{\#}$ is normal in $G^{\#}$ and the dot denotes the semi-direct product. Let ξ_1 denotes the homomorphism of $G^{\#}$ onto $T^{\#}$ with kernel $N^{\#}$. Then if ad denotes the adjoint representation of G, we have

$$\xi_1 \,\mathrm{ad}\colon G \to T^{\#}$$

is a homomorphism of G into $T^{\#}$. Let $\xi = \xi_1 \,\mathrm{ad}$ *be called the semi-simple homomorphism and let* $T_S = \xi(G)$.

It is crucial for our purposes that we can consider T_S as a group of automorphisms of G. Now from Chevalley [6] Chapter IV, § XV, we see that the group of automorphisms of the Lie algebra of G, which certainly contains $\mathrm{ad}(G)$, is an algebraic group. Hence $G^{\#}$ is a group of automorphisms of G, since G is simply connected. Since we may consider T_S as a subgroup of $G^{\#}$, we may consider T_S as a group of automorphisms of G. We now form the group $G_S = T_S \cdot G$, where the dot as usual denotes the semi-direct product.

PROPOSITION 4.1. *Let* $G_S = T_S \cdot G$ *be defined as above and let* N_S *denote the nil-radical of* G_S. *Then*

1) $G_S = T_S \cdot N_S$,

2) *the nil-radical* N *of* G *is normal in* G_S *and* G_S/N *is abelian.*

Proof. Since G is normal in G_S and N is characteristic in G, N is normal in G_S. Hence $N_S \supset N$. Further, we have obviously that $T_S \cap N_S$ is the identity. Now noting that T_S acts trivially on G/N, we may conclude that G_S/N is abelian.

Now let us consider $\mathrm{ad}(G_S)$ restricted to N. Then, as the discussion preceding the statement of our Proposition, we determine a homomorphism ψ of G_S onto a semi-simple abelian group of automorphisms of N. This group is clearly isomorphic to T_S. Let K denote the kernel of ψ. Then K clearly contains N_S. Since K is normal it remains only to verify that K is connected and nilpotent. But since G_S/N is abelian, and K/N acts unipotently on N, we have that K is nilpotent. We may easily verify that $G_S = T_S \cdot K$. Hence K is homeomorphic to G and therefore connected and simply connected. Hence $K = N_S$.

THEOREM 4.1. *Let G be a solvable Lie group. Then there exists a unique nilpotent Lie group N_S and a unique abelian semi-simple group T_S of automorphisms of N_S such that*

a) $G \subset T_S \cdot N_S$.

b) *Let $p\colon T_S \cdot N_S \to N_S$ be the projection mapping onto the second factor. Then p restricted to G is a homeomorphism of G onto N_S.*

c) *G and N_S generate $T_S \cdot N_S$.*

We have already verified the existence part of the theorem in Proposition 1 and hence it remains only to prove the uniqueness part of the assertion. The proof of this theorem will be presented as a sequence of remarks and lemmas.

Remark. 1. Let A be a semi-simple group of automorphisms of a nilpotent Lie group N which is trivial on $N/[N,N]$. Then A acts trivially on N. We can see this easily as follows: $[N,N]$ is an invariant subspace for A and hence there exists a linear subspace W of N such that W is invariant under A and $N = W \oplus [N,N]$ as a vector space. Further, by our hypothesis A acts trivially on W. But, since N is nilpotent, W generates N as a Lie algebra. Hence, since A was a group of automorphisms A acts trivially on N.

Remark 2. Let $S = T_S \cdot N_S$. Then $[S,S] \subset N_S$ and hence $[G,G] \subset N_S$.

Remark 3. Let $\psi\colon N_S \to N_S/[N_S,N_S]$ be the natural homomorphism and let $M = \psi^{-1}(\psi((N_S \cap G)_0))'$ where $(N_S \cap G)_0$ denotes the identity component of $(N_S \cap G)$. Then M is a normal closed subgroup of S. We may see this last statement as follows: Any subgroup of N_S containing $[N_S,N_S]$ is normal in N_S and hence M is normal in N_S. But $(N_S \cap G)_0$ is normal in G and since $[N_S,N_S]$ is characteristic in N_S, $\psi^{-1}(\psi((N_S \cap G)_0)) = M$, must be invariant under inner automorphisms by elements of G. But, since G and N_S generate S, M must be normal in S. The fact that M is closed follows from the fact that is is a connected subgroup of a simply connected nilpotent subgroup.

Remark 4. Let $S/M = S'$ and let us consider the image of G in S', call it G'. Let p' be the projection of S' onto N_S/M. Then p' restricted to G' maps it onto N_S/M. Clearly, N_S/M is abelian. Now $G/(N_S \cap G)_0$ is abelian, since $[G,G] \subset [S,S] \subset N_S$ and $[G,G]$ is connected. Hence G' is abelian.

LEMMA 1. *S' is an abelian group.*

Proof. Let T' be the quotient group of T_S which acts faithfully on N_S/M. We will show that T' is trivial which is equivalent to S' being abelian. Let $S'' = T'.\ (N_S/M)$ and let $(\eta): S' \to S''$ be the natural homomorphism. By identifying N_S/M with its Lie algebra, we can consider it as a real vector space and we have that T' acts as a group of automorphisms of this vector space. Hence S'' may be considered as a subgroup of the group of affinities of the affine space N_S/M. Since $\eta(G')$ and N_S/M generate S'', if T' is non-trivial, there exists $g \in \eta(G')$ whose T' component is non-trivial.

Now let W be the subspace of N_S/M spanned by elements from $(G/M \cap G) \cap (N_S/M)$. Since G' is abelian, W must be an eigen value one space for T'. This is straightforward to verify. Hence W is normal in S''. We could divide by it and proceed or equivalently, we will just assume it is trivial. This will avoid the necessity of introducing more notation. Then G' is an abelian group of affinities which operates transitively on the affine space .. It follows from results in [4a] and other places, that G' has all eigen values equal to one. This shows that T' is trivial and S' is abelian.

LEMMA 2. *If $G \supset [N_S, N_S]$, then $S/[G, G]$ is abelian.*

Proof. $[G, G] \subset N_S$ and is connected. Further, G is normal in S by Lemma 1, since S/M is abelian and $M = (N_S \cap G)_0$, by our hypothesis. Hence $[G, G]$ being characteristic in G is normal in S. Hence T_S leaves $[G, G]$ invariant and induces a group of automorphisms on $N_S/[G, G]$. But T_S is determined by its action on $N_S/[N_S, N_S]$. Hence T_S on $N_S/[G, G]$ is determined by its action on N_S/M. But by Lemma 1 this is trivial. Hence T_S induces a trivial action on $N_S/[G, G]$. Hence the projection of $G/[G, G]$ on $N_S/[G, G]$ is an isomorphism and $[G, G] \supset [N_S, N_S]$. Hence $S/[G, G]$ is abelian.

LEMMA 3. $G \supset [N_S, N_S]$.

Proof. We will carry out the proof by induction on the dimension of G. There will be two special cases to be treated. Case 1: Assume there is a non-trivial connected subgroup H in $G \cap [N_S, N_S]$ which is normal in S. Then let $S' = S/H$, $G' = G/H$, $N' = N_S/H$ and let T' equal T_S restricted to N'. Then the restriction mapping of T_S to T' is an isomorphism by Remark 1. Clearly S', N', T' and G' satisfy our inductive hypothesis. Further $g' \in G'$ is in M' if and only if its pre-image in G is in N_S. Hence $G' \supset [N', N']$ implies easily that $G \supset [N_S, N_S]$.

Case 2. The only connected subgroup in $G \cap [N_S, N_S]$ which is normal in S has dimension zero. Note first that if N^k is the last non-trivial term of the lower central series of N_S then $(G \cap N^k)_0$ is trivial. Otherwise it would satisfy the hypothesis of Case 1.

Let $g_0 \in G$ be such that $p(g_0) \in N^k$. Then $g_0 = (t_0, n_0)$, $t_0 \neq 0$. Now $(N_S \cap G)_0$ is invariant under $\mathrm{ad}(g_0)$, which since n_0 is central in N_S, is the same as $\mathrm{ad}(t_0)$. We will next verify that

$$(\mathrm{ad}(t_0) - I)N_S \subset (N_S \cap G)_0$$

where we have in this section identified N_S with its Lie algebra, by means of the exponential mapping. To see this, let $g = (t, n)$, where n is an eigen vector for $\mathrm{ad}(t_0)$. Then if $[\ ,\ ]$ denotes the commutator operation

$$[g_0, g] = (\mathrm{ad}(t_0) - I)n + (1 - \mathrm{ad}(t))n_0 \subset (N_S \cap G)_0$$

But $(\mathrm{ad}(t) - I)n_0 \in \mathrm{N}^k$. Hence

$$[g_0, [g_0, g]] = (\mathrm{ad}(t_0) - I)^2 n$$

since N^k must be a null space for $(\mathrm{ad}(t_0) - I)$ for otherwise $(N^k \cap G)_0$ would not be non-empty. We may see this as follows: Let a new semi-direct product representation of S be chosen so that n_0 is in the eigen value 1 space for $\mathrm{ad}(t_0)$. Let $g = (t_1, n_1)$ have the property that n_1 is an element of the range of $(\mathrm{ad}(t_0) - I)$. Then

$$[g_0, g] = (\mathrm{ad}(t_0) - I)n_1 + (I - ad(t))n_0$$

Since $\mathrm{ad}(t_0)$ and $\mathrm{ad}(t)$ commute the range of $(\mathrm{ad}(t_0) - I)$ is invariant under $\mathrm{ad}(t)$ as is the eigen-value one space of $\mathrm{ad}(t_0)$. Hence

$$(\mathrm{ad}(t_0) - I)n_1 + (I - \mathrm{ad}(t))n_0$$

is not empty. Hence if n is an eigen vector, with eigen value not equal to one, then $n \in (N_S \cap G)_0$. The argument for two dimensional invariant spaces is then essentially the same. Hence, since $\mathrm{ad}(t_0)$ is semi-simple, we have

$$(\mathrm{ad}(t_0) - I)N_S \subset (N_S \cap G)_0.$$

But $(\mathrm{ad}(t_0) - I)N_S$ is invariant under T_S and hence we have a vector space $W \subset (N_S \cap G)_0$ invariant under T_S and such that $W \cap [N_S, N_S]$ is trivial. Now W is non-empty, since otherwise by Remark 1, $\mathrm{ad}(t_0)$ would be trivial acting on $N/s[N_S, N_S]$ and hence on N_S. Let $\{W\}_1$ be the ideal generated by W in G and $\{W\}_2$ the ideal generated by W in N_S. Then $\{W\}_1 = \{W\}_2$ and we will henceforth call this space $\{W\}$. For

$$\{W\}_2 = W + \Sigma[n_i, w_i] n_i \in N_S \text{ and } w_i \in W$$

But

$$\{W\}_1 = W + \Sigma[(t_i, n_i), W_i] = W + \Sigma[n_i, w_i]$$

since W is invariant under T_S. But $\{W\} \cap [N_S, N_S]$ must be trivial otherwise we would violate the hypothesis of Case 2. Thus $W = \{W\}$ and W is central in N_S.

Again form $N' = N_S/W$, $G' = G/W$ and T' the restriction of T_S to N_S/W. This time the restriction map of T_S to T' will not be an isomorphism. Hence we cannot argue again as in Case 1. But we now employ Lemma 2. Noting that N', G', T' satisfy our inductive hypothesis we get the image g'_0 of g_0 is in $[N', N']$ and hence we have g'_1 and $g'_2 \in G'$ such that $[g'_1, g'_2] = g'_0$. This leads to the result that $n_0 \in (G \cap N_S)_0 \operatorname{mod} W$, since W is central. But since $W \subset (G \cap N_S)_0$, $n_0 \in (G \cap N_S)_0$. This contradicts our hypothesis that $(N^k \cap G)_0$ is empty and Case 2 has been completely treated and our lemma is proven.

Proof of Theorem 4.1. By Lemma 2 $G \supseteqq [S, S]$ and hence G is normal in S. Further, M is clearly in the nil-radical of G and since $G \cap T_S$ is trivial we must have also that $S = T_S \cdot G$. It is then straight forward to verify that $T_S = \xi(G)$ as defined in the beginning of this section and that N_S is the nil-radical of S. This proves our assertion.

In view of this result, we can make the following definition.

Definition. For each G we will call $T_S \cdot N_S$ as discussed in the previous theorem the semi-simple splitting of G and denote it by G_S.

If G is type (R) then T_S is a subgroup of the orthogonal group, but it need not be closed. We will use T^* as the closure of T_S in this group. Let $G_S^* = T^* \cdot G \supset G_S$.

THEOREM 4.2. *Let C be a closed subgroup of G such that G/C is compact. If G is type R and $G_S = T_S \cdot N_S$ is the semi-simple splitting of G then T_S is compact and*

$$C/C \cap N_S \text{ is finite.}$$

Proof. Clearly G_S^*/C is compact. Let G_1^* be the universal covering group of G_S^* and let C^* be the pre-image of C in G_1^*. Then $G_1^* = V \cdot N_S$, where V is the vector space covering group of T^* and N_S is the nil-radical of G_1^*. Then by the main theorem of Mostow [12] $N_S/C^* \cap N_S$ is compact and the identity component of C^*, C_0^*, is contained in $C^* \cap N_S$. Further the image of C^* in $G_1^*/N_S = V$ is a lattice. Hence $C/C \cap N_S$ must be finite.

Consider N_S acting on G_S/C. This is a fiber bundle over $T_S(C/C \cap N_S)$ with fiber $N_S/C \cap N_S$. Now S/C maps onto $T_S/(C/C \cap N_S)$ and hence, since the continuous image of a compact space is compact, $T_S/(C/C \cap N_S)$ is compact. Since $C/C \cap N_S$ is finite T_S is compact and hence $T_S{}^* = T_S$.

Corollary 4.1. *Let G be a solvable Lie group of type (R) and let C be a closed subgroup of G such that G/C is compact. Then G/C is finitely covered by the nilmanifold, $N_S/C^* \cap N_S$.*

This corollary follows from the uniqueness theorem for solvmanifolds; i.e., compact solvmanifolds with isomorphic fundamental groups are homeomorphic.

Definition. Let $G_S = T_S \cdot N_S$ where T_S is a compact abelian group of automorphisms of N_S and N_S is a simply connected, connected, nilpotent Lie group. Let $\phi(\tau) = (t(\tau), n(\tau))$ be a one parameter group in G_S. Then we say that $\phi(\tau)$ is in general position if $t(\tau)$ is dense in T_S.

Theorem 4.3. *Let G be a solvable Lie group with closed subgroup C such that G/C is compact and semi-simple splitting $G_S = T_S \cdot N_S$. Further, let $\phi(\tau)$ be a one parameter subgroup of G. If the flow $(G/C, \phi)$ has a dense orbit then $\phi(\tau)$ is in general position.*

This is easily verified.

Remark. This condition is always satisfied by our induced flows since they come from really regular elements.

Theorem 4.4. *Let G be a solvable Lie group with semi-simple splitting $G_S = T_S \cdot N_S$. Further let C be a closed subgroup of G such that*

1. *G/C is compact,*
2. *$C \cap N_S = C$.*

Then the flow of a one parameter group in general position in G/C is equivalent to a flow induced by a one parameter group of N_S/C.

The proof of this is identical to the proof of the Theorem 4.2, page 71 of [1]. By the results on nilflows in [1] we have the following conclusion.

Corollary 4.2. *Assume the hypothesis of the above theorem are satisfied. Then the following statements are equivalent.*

1) *$(G/C, \phi)$ has one dense orbit,*

2) *$(G/C, \phi)$ is minimal,*

3) *$(G/C, \phi)$ is ergodic.*

THEOREM 4.5. *Let G be a solvable group of type (R) and $G_S = T_S \cdot N_S$ be its semi-simple splitting. Then the flow induced by a one parameter group in general position in G is ergodic on G/C if and only if it is ergodic on $G/C \cap N_S$.*

Proof. It is trivial to verify that the flow on $G/C \cap N_S$ is distal and hence the flow on G/C is distal by Theorem 5.1 of Chapter III of [1]. This means that if the flow on G/C is ergodic, it is minimal, since ergodic implies minimal when the flow is distal. Hence by Theorem 5.1 of Chapter III of [1] the flow is minimal on $G/C \cap N_S$ and by the previous corollary it is ergodic. But if it is ergodic on $G/C \cap N_S$ it is ergodic on G/C.

Appendix I. We have seen that if G is a type (R) solvable Lie group and C is a co-compact subgroup, then it is finitely covered by a nilmanifold. We will in this appendix establish the converse of this result. Our main result may be stated as follows:

THEOREM I.1. *Let G/C be a compact solvmanifold with a nilpotent fundamental group. Then G is of type (R).*

The proof of this main theorem will follow easily once we have established the following results.

In the proof of the following theorem we will use freely results on algebraic groups for which Borel [5] will serve as a general reference. We will also use facts on solvmanifolds which are established in Mostow [12].

THEOREM I.2. *Let G be a connected simply connected solvable Lie group with C as a co-compact subgroup such that the fundamental group of $G/C, \pi_1(G/C)$, is a nilpotent group. Then C is nilpotent.*

Proof. Let $\mathrm{ad}_0(C)$ denote the automorphism group of $L(C_0)$, the Lie algebra of C_0, induced by inner automorphisms of C. Then, since C_0 and C/C_0 are nilpotent, C itself is nilpotent if and only if each linear transformation in $\mathrm{ad}_0(C)$ has all of its eigenvalues equal to one.

Since G is a connected simply connected solvable Lie group, G has a faithful matrix representation. Let $G^\#$ denote the algebraic hull of G in this matrix representation. Then $G^\# = T \cdot U$, where U is the normal subgroup of $G^\#$ consisting of unipotent matrices and T is a semi-simple abelian group. Let N be the maximal analytic normal nilpotent subgroup of G.

We may, by choosing our original representation appropriately, assume that $N \subset U$. Let us now consider the group CN and its algebraic hull G^*. Then $G^* \subset G^\#$, since $CN \subset G$. Further, if $G_0{}^*$ denotes the identity component of G^*, then $G_0{}^* = T^* \cdot U$ for $T^* \subset T$. This may be seen as follows: $CN = \gamma_1{}^{i_1} \cdot \cdot \cdot \gamma_k{}^{i_k} N$, where the image of the $\gamma_1, \cdot \cdot \cdot, \gamma_k$ in G/N form a basis for this vector space. Hence if $\gamma_i = u_i t_i$ and where $u_i t_i = t_i u_i$ and u_i is unipotent, we have $u_i \in U$ all i and the unipotent subgroup of $G_0{}^*$ is the algebraic hull of $\{u_i\}$ and N. This is clearly U. This shows that $G_0{}^* = T^* \cdot U$.

Now C_0 is normal in CN, since it is normal in both C and N. Hence, since any normal connected subgroup of a matrix group is normal in the algebraic hull of the matrix group, C_0 is normal in G^*. Since $C_0 \subset U$ and $U \subset G^*$, we have that C_0 is normal in U. Let $\mathrm{ad}_U(G^\#)$ denote the adjoint representation of $G^\#$ restricted to the Lie algebra of U, $L(U)$. Let the image of CN in $\mathrm{ad}_U(G^\#)$ be denoted by D. In order to prove that C is nilpotent it will suffice to show that the matrices of D have all their eigenvalues equal to one.

We first note that NC/C_0 is nilpotent. This is because C_0 is normal in N and $C/C_0 \cap N/C_0$ is a discrete uniform subgroup of N/C_0. Since C/C_0 is nilpotent it induces unipotent automorphisms of C/C_0 and hence of N/C_0 by the unique extension theorem of Malcev (Theorem 5, [10]). Let K_1 be the subspace of $L(U)$ which is the range of $(t - I)$ for $d \in D$, and $d = t \cdot n_1$. This must be in C_0, otherwise NC/C_0 would not be nilpotent. Since T is abelian K_1 is invariant under T. Then $K_1 \subset L(C_0) \subset L(U)$. Let $I \subset L(U)$ be the ideal generated by K_1 in $L(U)$. Since T acts as automorphisms of U, the ideal in $L(U)$ generated by K_1 must be invariant under T. This is because I is composed of sum of the form $\Sigma[k_i, X_i] + K_1$, where $k_i \in K_1$. But for $t \in T$, $t\Sigma[k_i, X_i] = \Sigma[k_i', X_i']$ and $tK_1 \subset K_1$, $k_i' \in K_1$. This proves that I is invariant under T. But since $K_1 \subset L(C_0)$ and C_0 is normal in U, we have $I \subset L(C_0)$. Thus the subgroup of H of $G^\#$ with Lie algebra I is a normal subgroup of $G^\#$. We see that $H \subset C_0$ and H is normal in G.

Since C_0 contains no analytic normal subgroup of G, it follows that $H = \{e\}$ and thus $I = \{o\}$. Thus $K_1 = \{o\}$ and $d \in D$ has all eigenvalues equal to one.

THEOREM I. 3. *Let G be a nilpotent Lie group with identity component G_0. If*

1) *G_0 is simply connected,*

2) *G/G_0 is finitely generated and torsion free, then there is a unique connected simply connected nilpotent Lie group $N(G) \supset G$ for which $N(G)/G$*

is compact. Moreover each automorphism of G extends uniquely to an automorphism of $N(G)$.

The proof of this theorem is essentially the same as the proof of Theorem 6 of Malcev [10] and we refer the reader to Malcev [10].

THEOREM I.4. *Let $\pi_1(G/C)$ be nilpotent. There is a connected simply connected nilpotent Lie group $N(C) \supset C$ for which $N(C)/C$ is compact and C_0 is normal in $N(C)$. There also is a torus group K of automorphisms of $N(C)$ such that $G \subset N(C) \cdot K$ and the projection π of $N(C) \cdot K$ onto $N(C)$ induces a homeomorphism of G/C onto $N(C)/C$.*

Proof. We first observe that it follows from Theorem I.2 that C is nilpotent. Since $C_0 \subset N$ we conclude that C_0 is simply connected. By applying Theorem 2, we may conclude that there is a unique connected, simply connected, nilpotent Lie group $N(C) \supset C$ for which $N(C)/C$ is compact. The theorem is now concluded in the same fashion as Theorem 1 in [4]. The argument is unchanged.

Theorem I.1 is a consequence of the fact that $G \subset N(C) \cdot K$ and K is compact.

Appendix II. In this section we give an elementary proof of what we called in [1] (Chapter II) "Mautner's Lemma." However, we do not obtain that detailed description of the spectrum which is used to prove mixing in the semisimple case.

Let S_1 be the affine group of the real line, represented as the set of real matrices of the form $\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$, $a > 0$. As a basis for its Lie algebra, choose $X = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $Y = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. For S_2 we consider the three-dimensional solvable group with Lie algebra basis $\{X, Y, Z\}$ and relations $[X, Z] = \sigma Y + Z$, $[X, Y] = Y - \sigma Z$, $[Y, Z] = 0$, where σ is a non-zero real number. Let H_i be the commutator group of S_i, $i = 1, 2$.

THEOREM. *Let $U\colon g \to U_g$ be a continuous unitary representation of S_i, $i = 1, 2$. If $U(\exp tX)\psi = e^{i\lambda t}\psi$ for some real constant λ and all t, then $U_h\psi = \psi$ for all $h \in H_i$.*

Proof. For either group in question, set $\phi_n(g) = \exp(-nX) g \exp(nX)$. By direct computation with matrices it is quickly verified that, in S_1,

$$\lim_{n\to\infty} \phi_n \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}.$$

Using the representation of S_2 by pairs (t, w), t real, w complex, with product

$$(t, w)(t', w') = (t + t', w + e^{(1-i\sigma)t}w'),$$

we compute that

$$\phi_n(t, w) = (-n, 0)(t, w)(n, 0) = (t, e^{-n}e^{i\sigma n}w).$$

Hence for each of these groups,

$$\lim_{n\to\infty} \phi_n(h) = e$$

for every $h \in H_i$.

Now suppose ψ satisfies the hypotheses of the Theorem, and fix $h \in H_i$. Then

$$\begin{aligned}(U_h\psi, \psi) &= (U(\exp nX)U((\exp - nX)h(\exp nX))U(\exp - nX)\psi, \psi) \\ &= (U(\phi_n(h))e^{-in\lambda}\psi, U(\exp - nX)\psi) \\ &= (U(\phi_n(h))\psi, \psi) \to (\psi, \psi),\end{aligned}$$

since U is a continuous representation. Hence $(U_h\psi, \psi) = (\psi, \psi)$, and (using the Schwarz inequality and the fact that U is unitary), $U_h\psi = \psi$.

The idea of this proof appears in Segal and von Neumann, "A theorem on unitary representations of semi-simple Lie groups," *Annals of Mathematics*, vol. 52 (1950), pp. 509-517. Clearly Theorem 2.4 may also be reformulated to give information about the common kernel of all representations into finite factors (of an arbitrary Lie group) by using their Lemma 1 and the decomposition given in our Lemma 2.3.

REFERENCES.

[1] L. Auslander, L. Green and F. Hahn, *Flows on homogeneous spaces*, Annals of Mathematical Studies No. 53, Princeton University Press, Princeton, N. J., 1963.

[2] L. Auslander and L. Green, "Flows on solvmanifolds," *Bulletin of the American Mathematical Society*, vol. 69 (1963), pp. 745-746.

[3] L. Auslander and F. Hahn, "Discrete transformations on tori and flows on solvmanifolds," *Bulletin of the American Mathematical Society*, vol. 68 (1962), pp. 614-615.

[4] L. Auslander, "Solvable Lie groups acting on nilmanifolds," *American Journal of Mathematics*, vol. 82 (1960), pp. 653-660.

[4a] ———, "On the group of affinities of locally affine spaces," *Proceedings of the American Mathematical Society*, vol. 9 (1958), pp. 471-473.

[5] A. Borel, "Groupes linéaires algébriques," *Annals of Mathematics*, vol. 64 (1956), pp. 20-82.

[6] C. Chevalley, *Theory of Lie Groups*: Vol. I," Princeton University Press, 1946.

[6a] ———, *Théorie des Groupes de Lie*: Tome III, Hermann, Paris, 1955.

[7] H. Furstenberg, "The structure of distal flows," *American Journal of Mathematics*, vol. 85 (1963), pp. 477-515.

[8] F. J. Hahn, "On affine transformations of compact abelian groups," *American Journal of Mathematics*, vol. 85 (1963), pp. 428-446.

[9] G. W. Mackey, "Borel structure in groups and their duals," *Transactions of the American Mathematical Society*, vol. 85 (1957), pp. 134-165.

[9a] ———, "Ergodic transformation groups with a pure point spectrum," *Illinois Journal of Mathematics*, vol. 8 (1964), pp. 593-600.

[10] A. Malcev, *On a class of homogeneous spaces*, American Mathematical Society Translations, vol. 39 (1951).

[11] ———, *Solvable Lie algebras*, American Mathematical Society Translations, vol. 27 (1950).

[12] G. D. Mostow, "Factor spaces of solvable groups," *Annals of Mathematics*, vol. 60 (1954), p. 127.

[13] ———, "Homogeneous spaces with finite invariant measure," *Annals of Mathematics*, vol. 75 (1962), pp. 17-37.

ON TWO PROBLEMS OF MORDELL.

By E. Bombieri and H. Davenport.

1. In recent papers, Mordell [6], [7] has propounded two problems which relate to the finite field $[p]$ of p elements, where p is a large prime. In the first, $f(x)$ is a polynomial over $[p]$ of fixed degree $d \geqq 2$, and the problem is whether we can assert the existence of some integer m, relatively small compared with p, for which the congruence

$$f(x) \equiv m \pmod{p} \tag{1}$$

is insoluble. Obviously we must exclude the possibility that $f(x)$ is a substitution polynomial $(\operatorname{mod} p)$, since then the congruence has just one solution for each m. To the best of our knowledge, no simple explicit condition is known which is both necessary and sufficient for a polynomial to be a substitution polynomial $(\operatorname{mod} p)$. But it is known (see Davenport and Lewis [3]) that if the polynomial

$$\phi(x,y) = (f(x) - f(y))/(x - y) \tag{2}$$

has the property that at least one of its irreducible factors $(\operatorname{mod} p)$ is absolutely irreducible, then $f(x)$ is not a substitution polynomial. We shall prove:

Theorem 1. *Let $f(x)$ be a polynomial over $[p]$ of degree $d \geqq 2$, and suppose that at least one of the irreducible factors of $\phi(x,y)$ over $[p]$ is absolutely irreducible. Then there exists a number $C(d)$ depending only on d such that, for every large p, there is some m with*

$$0 < m < C(d) p^{\frac{1}{2}} \log p \tag{3}$$

for which the congruence (1) *is insoluble.*

The bound for m is almost the same as that which was found by Mordell in the case $d = 3$.

The second problem relates to an exponential sum in two variables. Let

$$F(x,y) = f_3(x,y) + f_2(x,y) + f_1(x,y) \tag{4}$$

be a cubic polynomial over $[p]$, which we suppose not to be degenerate (that

---

Revised received February 24, 1965.

is, not to be expressible after a linear change of variables as a polynomial in one variable). The problem is that of estimating

$$(5) \qquad S_1(F) = \sum_{x=0}^{p-1} \sum_{y=0}^{p-1} e_p(F(x,y)), \text{ where } e_p(z) = e^{2\pi i z/p}.$$

It was proved by Davenport and Lewis ([2], Theorem 2) that

$$(6) \qquad S_1(F) = O(p^{5/4}),$$

and Mordell proved that for certain types of cubic polynomial

$$(7) \qquad S_1(F) = O(p).$$

We shall prove, in the following theorem, that this holds generally.

THEOREM 2. *Let $F(x,y)$ be a non-degenerate cubic polynomial over $[p]$. Then*[1]

$$(8) \qquad |S_1(F)| < 14p.$$

The proofs of both theorems depend, in an essential manner, on the recent work of Bombieri [1] on the L-functions corresponding to multiple exponential sums, this work being based in turn on that of Dwork [4, 5] on the congruence zeta-functions corresponding to algebraic varieties. The essential difficulty in Theorem 1 is that of estimating an exponential sum 'along a curve,' that is (in the present case) a sum of the type

$$\sum_{\substack{x,y \\ \phi(x,y)\equiv 0 \pmod p}} e_p(f(x)).$$

The essential difficulty in Theorem 2 lies in relating the sum (5) to the analogous sums over finite fields $[p^\nu]$ for which ν is a multiple of 3. These latter sums are easily estimated by appealing to Weil's results on exponential sums formed with rational functions of one variable. In the course of the argument we make use of the imperfect estimate (6).

2. The three lemmas which follow are needed for the proof of Theorem 2; Lemma 4 gives the essential result needed for Theorem 1.

LEMMA 1. *Let $q = p^\mu$, where μ is a positive integer. Let $F(x_1, \cdots, x_n)$ be a polynomial over $[q]$ of degree $d \geqq 1$, and let*

$$(9) \qquad S_\nu^*(F) = \sum_{[q^\nu]}^{*} e_p(\sigma F(x_1, \cdots, x_n)),$$

---

[1] The constant 14 has no special significance, being merely named for convenience. The constant could be improved to 4, which is best possible, at the cost of considerable complications.

where $\sum^*$ *indicates that each variable runs through all elements except* 0 *of* $[q^\nu]$, *and where* σ *denotes the trace of an element of* $[q^\nu]=[p^{\mu\nu}]$ *relative to* $[p]$. *Then there exist algebraic integers* $\omega_1,\cdots,\omega_r,\eta_1,\cdots,\eta_s$, *independent of* ν, *with* $\omega_h\neq\eta_j$ *for every* h, j, *such that*

$$S_\nu^*(F)=(-1)^n(\omega_1{}^\nu+\cdots+\omega_r{}^\nu-\eta_1{}^\nu-\cdots-\eta_s{}^\nu) \tag{10}$$

for all ν. *Also*

$$0\leqq r-s\leqq d^n. \tag{11}$$

Further, if $s>0$, *we have*

$$\eta_j=q^{a_j}\omega_j,\ \textit{for}\ j=1,\cdots,s, \tag{12}$$

where a_j *is a non-zero integer; and if* $a_j<0$ *there exists* j' *such that*

$$\eta_j=\eta_{j'}\quad \textit{and}\quad a_{j'}>0. \tag{13}$$

Proof. This is Theorem 1 of Bombieri [1].

LEMMA 2. *Let* $F(x)$ *be a polynomial in one variable over* $[q]$ *of degree* d, *where* $(d,p)=1$, *and let*

$$S_\nu(F)=\sum_{[q^\nu]}e_p(\sigma F(x)). \tag{14}$$

Then there exist complex numbers $\theta_1,\cdots,\theta_{d-1}$, *independent of* ν, *such that*

$$S_\nu(F)=-\theta_1{}^\nu-\cdots-\theta_{d-1}{}^\nu \tag{15}$$

for all ν, *and these numbers satisfy*

$$|\theta_j|=q^{\frac{1}{2}}\qquad\qquad (1\leqq j\leqq d-1). \tag{16}$$

If, more generally, $F(x)$ *is a rational function whose poles (including possibly* $x=\infty$*) have multiplicity not divisible by* p, *and the sum* (14) *is extended over those* x *in* $[q^\nu]$, *and possibly* $x=\infty$, *for which* $F(x)$ *exists, then a result analogous to* (15) *holds, but with* $d-1$ *replaced by* $D=t+\deg(F)_\infty-2$, *where* t *is the number of distinct poles of* $F(x)$, *and* $\deg(F)_\infty$ *is the total number of poles of* $F(x)$, *counted with multiplicity (including* $x=\infty$ *among the poles if necessary).*

Proof. This is a special case of Theorem 5 of [1]; in this case the curve X in that theorem is the projective line, which is of genus 0.

LEMMA 3. *Let* b_m, z_m *be complex numbers,* $1\leqq m\leqq M$, *and* $|z_m|=1$, $z_r\neq z_s$ *for* $r\neq s$. *Then*

$$\limsup_{\nu\to+\infty}\Big|\sum_m b_m z_m{}^\nu\Big|\geqq\Big(\sum_m|b_m|^2\Big)^{\frac{1}{2}}. \tag{17}$$

Proof. The proof is based on an idea of P. Turán which uses a discrete analogue of the classical mean square integral in the theory of Fourier series. In fact, we have

$$\begin{aligned}\sum_{\nu=0}^{N} |\sum b_m z_m^{\nu}|^2 &= \sum_{\nu=0}^{N} \sum_{r,s} b_r \bar{b}_s (z_r/z_s)^{\nu} \\ &= (N+1)\sum |b_m|^2 + O(\sum_{r\neq s} |b_r b_s|\,|\sum_{\nu=0}^{N} (z_r/z_s)^{\nu}|) \\ &= (N+1)\sum |b_m|^2 + O(\sum_{r\neq s} |b_r b_s|/\|\arg(z_r/z_s)\|) \\ &= (N+1)\sum |b_m|^2 + O(1),\end{aligned}$$

because $z_r \neq z_s$. Letting $N \to +\infty$ we obtain the lemma.

LEMMA 4. *Let $f(x)$ and $\psi(x,y)$ be polynomials over $[p]$ such that $\deg(f) = d$, $1 \leqq d < p$, and $\psi(x,y)$ is absolutely irreducible and not of the type $\psi(x,y) = ax + b$, and has degree satisfying $1 \leqq \deg(\psi(x,y)) \leqq d$. Let*

$$S_\nu(f,\psi) = \sum_{\psi(x,y)=0} e_p(\sigma f(x)) \tag{18}$$

where the summation is over all $x, y \in [p^\nu]$ satisfying $\psi(x,y) = 0$. Then we have

$$|S_\nu(f,\psi)| \leqq 3((d^2-d)p^{\nu/2} + d^2. \tag{19}$$

Proof. We apply Theorem 6 of [1], where (in affine notation) X is given by $\psi(x,y) = 0$, $R(x) = f(x)$ and $d_1 = \deg(\psi)$, $d_2 = d$. In order to show that condition (A) of that theorem is satisfied, we have to prove that

$$f(x) \not\equiv h(x,y)^p - h(x,y) \pmod{\psi(x,y)},$$

for every rational function $h(x,y)$ defined over the algebraic closure K of $[p]$. We have $1 \leqq \deg f < p$, so that f is not of the form $f_1^p - f_1$ with f_1 in $K(x)$, as we can see by counting degrees in the equation $f = f_1^p - f_1$. It follows that if x is transcendental over K then the polynomial

$$z^p - z - f(x)$$

is irreducible over $K(x)$.

Let (x,y) be a generic point over K of the curve $\psi(x,y) = 0$, and suppose that

$$f(x) \equiv h(x,y)^p - h(x,y) \pmod{\psi(x,y)};$$

this means that the polynomial $z^p - z - f(x)$ has a root z in the function field $K(x,y)$ of the curve $\psi(x,y) = 0$. Hence the polynomial $z^p - z - f(x)$ is reducible over $K(x,y)$, and by a well-known property of Artin-Schreier extensions it follows that this polynomial splits completely over the field $K(x,y)$.

On the other hand, $\psi(x,y) \neq ax+b$ by hypothesis, whence if (x,y) is a generic point over K of the curve $\psi(x,y)=0$ then x must be transcendental over K. It follows that

$$[K(x,y):K(x)] \geqq p,$$

and this is impossible because

$$[K(x,y):K(x)] = \deg_y \psi(x,y) \leqq \deg \psi(x,y) < p.$$

This contradiction proves that Theorem 6 of [1] applies, and (19) follows.

3. Proof of Theorem 1. Let $r(m)$ denote the number of solutions of $f(x) \equiv m \pmod{p}$. Then

$$\sum_{m=1}^{p} r(m)\, e_p(mt) = \sum_{x=1}^{p} e_p(tf(x)) = S_1(tf(x)),$$

and

$$|S_1(tf(x))| \leqq (d-1)p^{\frac{1}{2}}$$

by Lemma 2 with $q=p$, provided $t \not\equiv 0 \pmod{p}$. By a well known argument, we have

$$\begin{aligned}\sum_{m=1}^{N} r(m) &= N + p^{-1}\sum_{t=1}^{p-1}\sum_{m=1}^{N} e_p(-tm)\, S_1(tf(x))\\ &= N + O(p^{-1}\sum_{t=1}^{p-1} p^{\frac{1}{2}} \max(p/t\,; p/(p-t)))\\ &= N + O(p^{\frac{1}{2}}\log p), \qquad (20)\end{aligned}$$

where the implied constant depends only on d.

Put

$$R(t) = \sum_{m=1}^{p} r^2(m)\, e_p(mt).$$

The same argument gives

$$\sum_{m=1}^{N} r^2(m) = \frac{N}{p} R(0) + O(R\log p),$$

where

$$R = \max|R(t)| \text{ for } t \not\equiv 0 \pmod{p}.$$

Now $R(0)$ is the number of solutions of

$$f(x) \equiv f(y) \pmod{p};$$

the number with $x \equiv y \pmod p$ is p, and the number with $x \not\equiv y \pmod p$ is the number of solutions of

$$\phi(x,y) \equiv 0 \pmod p, \quad x \not\equiv y \pmod p.$$

Omitting the last condition can only change the number by at most $d-1$, since $\deg \phi(x,x) \leqq d-1$. Let

$$\phi(x,y) = (\psi_1(x,y))^{u_1} \cdots (\psi_h(x,y))^{u_h},$$

where $\psi_1, \cdots, \psi_h$ are distinct and irreducible $(\operatorname{mod} p)$. Suppose $\psi_1, \cdots, \psi_e$ are absolutely irreducible, where $e \geqq 1$ by the hypothesis of Theorem 1, and suppose the others are not absolutely irreducible. By Weil's theorem, the number of solutions of $\psi_i(x,y) \equiv 0 \pmod p$ is $p + O(p^{\frac12})$ for $i \leqq e$. The number of solutions for $i > e$ is $O(1)$, as was proved in Davenport and Lewis [3], § 8. Hence

$$R(0) = (e+1)p + O(p^{\frac12}).$$

Also

$$R(t) = \sum_{f(x) \equiv f(y)} e_p(tf(x)) = \sum_x e_p(tf(x)) + \sum_{\substack{\phi(x,y) \equiv 0 \\ x \not\equiv y}} e_p(tf(x)).$$

The first sum is $O(p^{\frac12})$ for $t \not\equiv 0 \pmod p$. The second sum splits into sums over $\psi_1(x,y) \equiv 0$, etc., any two of which overlap by only $O(1)$ terms, and a residue of $O(1)$ terms for which $\phi(x,x) \equiv 0$. If $i \leqq e$, the sum over $\psi_i(x,y) \equiv 0$ is $O(p^{\frac12})$ by Lemma 4 with $\nu = 1$, and if $i > e$ the sum is $O(1)$. Hence

$$R = O(p^{\frac12}),$$

and it follows that

$$\sum_{m=1}^{N} r^2(m) = (e+1)N + O(p^{\frac12} \log p). \tag{21}$$

Let N_h denote the number of $m \leqq N$ for which $r(m) = h$, where $h = 0, 1, \cdots, d$. Then

$$N_0 + N_1 + \cdots + N_d = N,$$

and by (20) and (21)

$$N_1 + 2N_2 + \cdots + dN_d = N + O(p^{\frac12} \log p),$$
$$N_1 + 4N_2 + \cdots + d^2 N_d = (e+1)N + O(p^{\frac12} \log p).$$

Multiplying by $1, -1 - d^{-1}, d^{-1}$ respectively and adding, we deduce that

$$N_0 \geqq e d^{-1} N + O(p^{\frac12} \log p).$$

Thus $N_0 > 0$ if $N > C(d) p^{\frac{3}{4}} \log p$, for some suitable $C(d)$ (actually $C(d) = O(d^2)$). This proves the theorem.

4. Lemma 5. *Let $F(x,y)$ be a non-degenerate cubic polynomial over $[p]$, $p > 3$, and let*

$$S_\nu(F) = \sum e_p(\sigma F(x,y)), \tag{22}$$

where the summation is over all x, y in $[p^\nu]$. Then

$$S_{3\nu}(F) = \theta_1^{3\nu} + \cdots + \theta_D^{3\nu}, \tag{23}$$

where $D \leq 4$ and

$$|\theta_j| = p \tag{24}$$

for each j.

Proof. Let $f_3(x,y)$ denote the cubic part of $F(x,y)$, as in (4). There exist β, δ, not both zero, in $[p^3]$ such that $f_3(\beta,\delta) = 0$. By a linear transformation

$$x = \alpha x' + \beta y', \quad y = \gamma x' + \delta y' \qquad (\alpha\delta - \beta\gamma \neq 0)$$

we can express $F(x,y)$ as a cubic polynomial in x', y' with no term in y'^3. Such a transformation, with coefficients in $[p^3]$, is applicable if x, y are variables in $[p^{3\nu}]$, since $[p^3]$ is contained in $[p^{3\nu}]$.

Thus, without loss of generality, we can take

$$f_3(x,y) = ax^3 + bx^2y + cxy^2,$$

with a, b, c in $[p^3]$. Suppose first that $c \neq 0$. Then without loss of generality we can take $b = 0$, and by adding suitable constants to x and y we can further remove the terms in xy and y^2 in $F(x,y)$. Thus we can take

$$F(x,y) = ax^3 + cxy^2 + Ax^2 + Bx + Cy,$$

with coefficients in $[p^3]$.

Since the polynomial in y is quadratic, we can evaluate the sum over y in $[p^{3\nu}]$, using known results. If $p > 2$ and λ, μ are in $[p^3]$ and $\lambda \neq 0$, we have

$$\sum_{y \in [p^{3\nu}]} e_p(\sigma(\lambda y^2 + \mu y)) = -\theta^\nu \chi(\lambda) e_p(-\sigma(\mu^2/4\lambda)),$$

where θ is a Gaussian sum in $[p^3]$, with $|\theta| = p^{\frac{3}{2}}$, and χ is the quadratic character in $[p^{3\nu}]$. We have $\chi(\lambda) = \chi_1(\lambda)^\nu$, where χ_1 is the quadratic character in $[p^3]$. If $\lambda = 0$, the value of the sum is 0 if $\mu \neq 0$ and is $p^{3\nu}$ if $\mu = 0$.

On substituting, we obtain

$$S_{3\nu}(F) = -\theta^{\nu}\chi_1(\lambda)^{\nu}\sum_{x\neq 0}\chi(x)\,e_p(\sigma(ax^3+Ax^2+Bx-C^2/4cx)),$$

with an additional term $p^{3\nu}$ if $C=0$. The sum over x here is

$$\sum_{x\neq 0}\chi(x)\,e_p(\sigma(ax^3+Ax^2+Bx-C^2/4cx)) = -\sum_{h=1}^{D}\rho_h{}^{\nu}, \tag{25}$$

where $|\rho_h| = p^{\frac{3}{2}}$, by Weil's theorem on Artin L-functions (see (i) and (ii) of Theorem A of [1]).

This proves (23) and (24), and it remains to prove that $D \leqq 4$. The left hand side of (25) is

$$-\sum_{x\neq 0} e_p(\sigma(ax^3+Ax^2+Bx-C^2/4cx))$$
$$+\sum_{x\neq 0} e_p(\sigma(ax^6+Ax^4+Bx^2-C^2/4cx^2));$$

now by the second part of Lemma 2 we have if $p>3$:

$$\sum_{x\neq 0} e_p(\sigma(ax^3+Ax^2+Bx-C^2/4cx)) = -\sum_{h=1}^{4}\phi_h{}^{\nu},$$

and also

$$\sum_{x\neq 0} e_p(\sigma(ax^6+Ax^4+Bx^2-C^2/4cx^2)) = -\sum_{h=1}^{8}\psi_h{}^{\nu}.$$

Comparison of these two last equations with (25) shows that $D = 8-4 = 4$; in a similar way we handle the subcases $C=0$ and $a=0$.

There remains for consideration the case $c=0$, which gives rise to further subcases; but these can all be treated on the same lines, with the same result. We obtain $D=4$ if $ac\neq 0$, and $D\leqq 3$ if $ac=0$; it may be of interest to remark that $D=4$ if and only if the equation $f_3(x,1)=0$ has three *distinct* roots.

LEMMA 6. *With the notation of Lemma* 5, *we have*

$$S_{\nu}(F) = O(p^{5\nu/4}). \tag{26}$$

Proof. For $\nu=1$, this is contained in Theorem 2 of Davenport and Lewis [2], and the proof extends to general ν without any significant change.

5. Proof of Theorem 2. We can suppose that $p>3$. By Lemma 1 with $n=2$ and $\mu=1$ we have

$$S_{\nu}^{*}(F) = \omega_1{}^{\nu}+\omega_2{}^{\nu}+\cdots+\omega_r{}^{\nu}-\eta_1{}^{\nu}-\cdots-\eta_s{}^{\nu},$$

where

$$(27) \qquad 0 \leqq r - s \leqq 9.$$

Also the final clause of Lemma 1 applies to the numbers ω_j, η_k. The relation between $S_\nu(F)$ and $S_\nu^*(F)$ is

$$S_\nu(F) = S_\nu^*(F) + S_\nu(F(x,0)) + S_\nu(F(0,y)) - \{e_p(F(0,0))\}^\nu.$$

The sums in one variable on the right are expressible as stated in Lemma 4, with $d = 3$, provided that $\deg_x F(x,0) = 3$, $\deg_y F(0,y) = 3$, but it is obvious that we may ensure this by making a preliminary linear transformation over $[p]$ of the polynomial $F(x,y)$. Hence

$$(28) \qquad \begin{aligned} S_\nu(F) = \omega_1^\nu + \cdots + \omega_r^\nu - \eta_1^\nu - \cdots - \eta_s^\nu \\ - \zeta_1^\nu - \zeta_2^\nu - \zeta_3^\nu - \zeta_4^\nu - \epsilon^\nu, \end{aligned}$$

where $|\zeta_j| = p^{\frac{1}{2}}$ and $|\epsilon| = 1$.

It follows from Lemma 6, applied with ν arbitrarily large, that

$$(29) \qquad |\omega_j| \leqq p^{5/4}, \qquad |\eta_k| \leqq p^{5/4}$$

for all j, k. The deduction may be made most readily by appealing to Lemma 3.

We now appeal to Lemma 5, which tells us that if ν is a multiple of 3 then

$$(30) \qquad S_\nu(F) = \theta_1^\nu + \cdots + \theta_D^\nu$$

where $|\theta_j| = p$ and $D \leqq 4$. Thus if ν is a multiple of 3, the terms on the right of (28) must reduce by cancellation to give the terms on the right of (30).

We shall prove that $s = 0$ or 1. Let us suppose that $s \geqq 2$, and that (after rearrangement)

$$|\eta_1| \leqq |\eta_2| \leqq \cdots \leqq |\eta_s|.$$

By (12) we have $\eta_1 = p^{a_1}\omega_1$, where a_1 is a non-zero integer. We can suppose that $a_1 > 0$, for if not there exists j' such that (13) holds, and we can simply interchange the suffixes 1 and j'.

We now have

$$|\omega_1| = p^{-a_1} |\eta_1| \leqq p^{\frac{1}{4}}$$

by (29). Also

$$|\omega_1| \leqq p^{-1} |\eta_1| \leqq p^{-1} |\eta_k|$$

for all k. Hence the term ω_1^ν on the right of (28) cannot cancel with any of the terms η_k^ν, ζ_k^ν, and also cannot equal any term θ_k^ν on the right of (30). It follows that ω_1^ν must cancel ϵ^ν, so that $|\omega_1| = 1$. Hence $|\eta_1| = p^{a_1}$, whence $|\eta_1| = p$ by (29).

We now have $|\eta_j| \geqq p$ for all j, and $|\eta_j| \leqq p^{5/4}$ by (29). Now $\eta_2 = p^{a_2}\omega_2$ by (12), and it is impossible that $a_2 < 0$ by (29). Hence $a_2 > 0$, whence $|\omega_2| \leqq p^{\frac{1}{4}}$. But this is impossible, since ω_2^ν cannot cancel against any ζ_k^ν and cannot equal any θ_k^ν, and because ϵ^ν was used to cancel ω_1^ν.

We have now proved that $s = 0$ or 1. If $s = 0$ then each ω_j^ν must either cancel one of ζ_k^ν or ϵ^ν or equal some θ_k^ν, whence $|\omega_j| = 1$ or $p^{\frac{1}{2}}$ or p. It follows from (28) and (27) that

$$|S_\nu(F)| \leqq 13p^\nu,$$

which implies (8).

Now suppose that $s = 1$. The possibility that $a_1 < 0$ does not arise, and by the argument given earlier we have $|\eta_1| = p$, $|\omega_1| = 1$ and $|\omega_j| \leqq p$ for $j > 1$. Hence, by (28) and (27),

$$|S_\nu(F)| \leqq \sum_{j=1}^{10} |\omega_j|^\nu + |\eta_1|^\nu + 4p^{\nu/2} + 1 < 14p^\nu,$$

which implies (8). This completes the proof of Theorem 2.

REFERENCES.

[1] E. Bombieri, "On exponential sums in finite fields," *American Journal of Mathematics* (in course of publication).

[2] H. Davenport and D. J. Lewis, "Exponential sums in many variables," *American Journal of Mathematics*, vol. 84 (1962), pp. 649-655.

[3] ———, "Notes on congruences, I," *Quarterly Journal of Mathematics*, vol. 14 (1964), pp. 51-60.

[4] B. Dwork, "On the rationality of the zeta-function of an algebraic variety," *American Journal of Mathematics*, vol. 82 (1960), pp. 631-648.

[5] ———, "On the zeta-function of a hypersurface," *Publications Mathématiques*, No. 12, *Institut des Hautes Etudes Scientifiques*, Paris (1962).

[6] L. J. Mordell, "On residues and non-residues of a polynomial," *Journal of the London Mathematical Society*, vol. 38 (1963), pp. 451-453.

[7] ———, "On a cubic exponential sum in two variables," *Journal of the London Mathematical Society*, vol. 38 (1963), pp. 356-358.

ON EXPONENTIAL SUMS IN FINITE FIELDS.

By ENRICO BOMBIERI.

The purpose of this paper is to prove some results of a general nature on the structure of exponential sums formed with polynomials in many variables over finite fields. The proofs of these results depend heavily on Dwork's p-adic theory and on a simple analogue of the Jensen formula for the zeros of an entire function.

We shall prove also some results of a more special nature which are useful for application to number theory; in fact, this research was motivated by the need to prove some generalizations of the classical estimates of A. Weil for exponential sums in one variable.

I should like to express here my great indebtedness to H. Davenport and B. Dwork for many stimulating conversations during the preparation of this paper, and to the referee who found some errors and inaccuracies in the first version, and also suggested corrections and improvements.

1°. The sums $S_m(f)$ and $S_m{}^*(f)$. Let $f = f(x) = f(x_1, \cdots, x_n)$ denote a polynomial in n variables, defined over $[q]$, the finite field with $q = p^a$ elements, where the prime number p is the characteristic of the field. Let σ denote the absolute trace from $[q^m]$ to $[p]$, and consider the two exponential sums

$$S_m{}^*(f) = \sum{}^* e_p(\sigma f(x_1, \cdots, x_n)) \tag{1}$$

$$S_m(f) = \sum e_p(\sigma f(x_1, \cdots, x_n)) \tag{2}$$

where as usual $e_p(x)$ denotes the function $e^{2\pi i x/p}$, and where the sum $\sum^*$ in (1) runs over the n-fold product of the multiplicative group of $[q^m]$, while the sum $\sum$ in (2) runs over the n-fold product of the whole field $[q^m]$.

The consideration of the two different types of sums comes from the following: while the sum $S_m(f)$ is the natural exponential sum associated with f, the sum $S_m{}^*(f)$ is the natural object of study in Dwork's p-adic theory, and in fact it turns out that these last sums have properties which have no direct analogue for the sums $S_m(f)$. See for instance Theorem 1, (ii) and (iii). On the other hand, we may write $S_m(f)$ as a sum of a finite number

---

Received May 22, 1964. Revised January 25, 1965.

of sums $S_m{}^*(f)$, in the following way: let A denote any subset (including the empty set) of the numbers $(1,2,\cdots,n)$, let $n(A)$ denote the number of elements in A, and let f_A be the polynomial in $n-n(A)$ variables, obtained from f by putting $x_h=0$ for $h\in A$. Then we have

$$S_m(f)=\sum_A S_m{}^*(f_A) \tag{3}$$

$$S_m{}^*(f)=\sum_A(-1)^{n(A)}S_m(f_A) \tag{4}$$

as it is easily seen using a simple combinatorial principle. (See Dwork [1]).

The main theorem we shall prove in this paper is the following

THEOREM 1. (i) *There exist algebraic integers* ω_h, η_j *for* $1\leqq h\leqq r$, $1\leqq j\leqq s$, *depending on* f *but not on* m, *different from* 0 *and such that* $\omega_h\neq\eta_j$ *for every* h, j, *such that:*

$$S_m{}^*(f)=(-1)^n\Big(\sum_{h=1}^{r}\omega_h{}^m-\sum_{j=1}^{s}\eta_j{}^m\Big); \tag{5}$$

(ii) $s\leqq r$, *and if* $s\geqq 1$ *then every* η_j *is of the type*

$$\eta_j=q^{a_j}\omega_j \tag{6}$$

where the a_j *are non zero integers;*

(iii) *if* $a_j<0$, *there is some* j' *with* $\eta_j=\eta_{j'}$, *and* $a_{j'}>0$;

(iv) *we have the inequality*

$$0\leqq r-s\leqq d^n \tag{7}$$

where $d=\deg(f)$. *Moreover, the inequality* (7) *is a best possible inequality for every* n *and every* d *satisfying* $(d,p)=1$.

II°. The functions $L^*(t,f)$ and $L(t,f)$. Let $L^*(t,f)$ and $L(t,f)$ be the two functions defined by

$$L^*(t,f)=\exp\Big(\sum_{m=1}^{\infty}S_m{}^*(f)\,t^m/m\Big) \tag{8}$$

$$L(t,f)=\exp\Big(\sum_{m=1}^{\infty}S_m(f)\,t^m/m\Big) \tag{9}$$

for $|t|<q^{-n}$.

This definition of functions L^* and L is clearly related to the definition of the Zeta function of an algebraic variety over a finite field as given by A. Weil; actually, as we shall explain in the last section of this paper, the function $L(t,f)$ is an Artin L-function related to Artin-Schreier-Witt extensions of the affine space A^n.

Using (4) we have easily the equation

$$L^*(t,f)=\prod_A L(t,f_A)^{(-1)^{n(A)}}. \tag{10}$$

We shall need a simple lemma on the behaviour of this function L when the variable t is in a neighbourhood of the origin, the neighbourhood being defined by means of an archimedean valuation, namely:

LEMMA 1. *The coefficients A_m of the Taylor series for $L^*(t,f)$ are algebraic integers in the algebraic number field K generated by a primitive p-th root of unity. If $A_m^{(h)}$ denotes any conjugate of A_m in K, then*

$$|A_m^{(h)}|\leqq R^m \tag{11}$$

for some R depending only on f and q.

Proof. In the affine space A^n of points $(x_1,\cdots,x_n)$ defined over $[q]$, let us consider the prime 0-cycles defined over $[q]$, namely: given $x=(x_1,\cdots,x_n)\in A^n$ defined over a finite algebraic extension of the ground field $[q]$, let $[q^m]$ be its smallest field of definition; the prime 0-cycle associated with x is the m-uple $\mathfrak{p}$ of points

$$\mathfrak{p}=x+x^q+\cdots+x^{q^{m-1}} \tag{12}$$

that is x together with all its conjugates over $[q]$. The degree of $\mathfrak{p}$ is clearly

$$\deg(\mathfrak{p})=m \tag{13}$$

and also, from the fact that f is defined over $[q]$, whence

$$f(x)^q=f(x^q), \tag{14}$$

we have that the value of the trace $\sigma f(x)$ is the same for x and all its conjugates. Thus we may write in a obvious notation $\sigma f(\mathfrak{p})$ as the trace $\sigma f(x)$ for any $x\in\mathfrak{p}$.

Hence the well known Euler product formula for Artin L-functions gives

$$L(t,f)=\prod_{\mathfrak{p}}(1-e_p(\sigma f(\mathfrak{p}))t^{\deg(\mathfrak{p})})^{-1}; \tag{15}$$

and combining this with formula (10) we have

$$L^*(t,f)=\prod_{\mathfrak{p}}{}'(1-e_p(\sigma f(\mathfrak{p}))t^{\deg(\mathfrak{p})})^{-1}, \tag{16}$$

where the meaning of $\prod'$ is that the product runs over all prime cycles $\mathfrak{p}$ such that if $\mathfrak{p}=x+x^q+\cdots+x^{q^{m-1}}$ and $x=(x_1,\cdots,x_n)$ then $x_1x_2\cdots x_n$

$\neq 0$. Now it is clear that $e_p(\sigma f(\mathfrak{p})) \in K$ and this proves the first part of Lemma 1.

The second part of Lemma 1 is a trivial consequence of the fact that the product in (16) (and any one of its conjugates in K) is majorized by

$$\prod_{\mathfrak{p}} (1 - t^{\deg(\mathfrak{p})})^{-1} = \exp(\sum_{m=1}^{\infty} q^{nm} t^m/m).$$

III°. p-adic theory. Let Q_p denote the completion of the field of rational p-adic numbers, and let Ω be the completion of the algebraic closure of Q_p. Let Ω_1 be the complete, totally ramified extension of Q_p obtained by adjoining to Q_p a primitive p-th root of unity, and let K_a denote the complete unramified extension of Q_p of degree a obtained by adjoining to Q_p the $(q-1)$-th roots of unity, where $q = p^a$. Then if Ω_0 is the unramified extension of Ω_1 of degree a we have a diagram of fields

$$\begin{tikzcd}
 & \Omega_0 \arrow[dl, dash] \arrow[dr, dash] & \\
\Omega_1 & & K_a \\
 & Q_p \arrow[ul, dash] \arrow[ur, dash] &
\end{tikzcd}$$

and $\Omega_0 = \Omega_1 \times K_a$ is the composite of Ω_1 and K_a.

The residue class field of Ω_0 and K_a is $[q]$, while the residue class field of Ω_1 and Q_p is $[p]$. The Frobenius automorphism $x \to x^p$ in $[q]$ may be lifted to an automorphism τ of Ω_0 leaving fixed a primitive p-th root of unity. It is easily seen that τ acts trivially on Ω_1 and Q_p, and the action of τ on Ω_0 and K_a is described by

$$\tau(\epsilon) = \epsilon^p$$

where ϵ is a $(q-1)$-th root of unity in Ω_0.

We shall denote by 'ord' the additive valuation in Ω with the normalization $\operatorname{ord}(p) = 1$, and by '$\operatorname{ord}_q$' the additive valuation normalized so that $\operatorname{ord}_q(q) = 1$.

We are interested in the behaviour of $L^*(t, f)$ as a function in Ω, and for this we shall use Dwork's theory with some small modifications. Hence in order to show where we have to modify his theory, we shall give here a short account of this method.

Let

$$E(t) = \exp(\sum_{j=0}^{\infty} t^{p^j}/p^j)$$

be the Artin-Hasse hyperexponential function in Ω, and let γ denote a root of

$$\sum_{j=0}^{\infty} \gamma^{p^j}/p^j = 0 \tag{17}$$

satisfying

$$\operatorname{ord}(\gamma) = 1/(p-1), \quad \gamma \in \Omega_1. \tag{18}$$

Then, in the terminology of Dwork [1], [2] (see [2] Lemma 4.1) the function

$$\theta(t) = E(\gamma t) = \sum_{m=0}^{\infty} \lambda_m t^m \tag{19}$$

is a splitting function, and all its coefficients λ_m satisfy

$$\operatorname{ord}(\lambda_m) \geqq m/(p-1), \quad \lambda_m \in \Omega_1. \tag{20}$$

Let us denote by u, v, w n-uples $(u_1, \cdots, u_n), \cdots$ of non-negative rational integers, and let X^u be the monomial $X_1^{u_1} \cdots X_n^{u_n}$; lastly, by $|u|$ we shall denote

$$|u| = u_1 + \cdots + u_n. \tag{21}$$

Then for every real b we shall denote by $L(b)$ the infinite dimensional linear space of all power series $\sum A_u X^u$ satisfying the conditions

$$\operatorname{ord}(A_u) \geqq b\,|u|/(p-1) + O(1), \quad A_u \in \Omega. \tag{22}$$

This definition of the space $L(b)$ is a slight modification of Dwork's, because, as we are working with non-homogeneous polynomials f, we need a different convergence condition (22). Also we have chosen a different b for defining $L(b)$, in order to show clearly that for the applications of the theory to the function $L(t, f)$ the choice of b depends on the polynomial f. Hence our spaces $L(b)$ are essentially Dwork's spaces $L(db/(p-1))$.

Let ψ_p be the endomorphism of the Ω_0-linear space $L(b)$ given by

$$\psi_p(\sum A_u X^u) = \sum A_{pu} X^u, \tag{23}$$

and for every fixed $F = F(X) = \sum F_u X^u \in L(b)$ let F denote the endomorphism of the Ω_0-linear space $L(b)$ given by

$$F\colon \xi \to F(X)\xi \quad \text{for } \xi \in L(b).$$

We are interested in the composite endomorphism of ψ_p and F, which is denoted by $\psi_p \circ F$. If $\xi = \sum c_u X^u \in L(b)$ we have

$$(\psi_p \circ F)(\sum c_u X^u) = \sum_u (\sum_v F_{pu-v} c_v) X^u;$$

it follows that $\psi_p \circ F$ has a matrix representation $(\psi_p \circ F)$ in Ω_0 indexed by pairs (u, v) of n-uples:

$$(\psi_p \circ F) = (F_{pu-v}) \tag{24}$$

where we have to put $F_{pu-v}=0$ if the n-uple $pu-v$ has some negative element.

Now, in order to have a satisfactory theory, we must be able to define the characteristic equation and the trace of this endomorphism, and this is by no means obvious because our space $L(b)$ is infinite dimensional. The natural way of doing this is to restrict ourselves to the consideration of spaces $L(b)$ with $b>0$; in this case we have:

(i) let $F\in L(b)$ with $b>0$; then for every b' with $0<b'\leqq b$, $\psi_p\circ F$ is an endomorphism of the Ω_0-linear space $L(b')$, and the characteristic function $(t\in\Omega)$

$$\det(I-(\psi_p\circ F)t)$$

is well defined (that is independent of the choice of the matrix representation $(\psi_p\circ F)$) and is independent of b';

(ii) the trace $\operatorname{Tr}(\psi_p\circ F)^m$ of the m-th iterate of $\psi_p\circ F$ is well defined and independent of b'.

The proof of this is given in Dwork [1]. We must mention that the theory of endomorphisms in p-adic Banach spaces is now available in Serre [7]; the two statements made above concerning our case mean that $\psi_p\circ F$ is a completely continuous endomorphism of $L(b)$ for every $b>0$ in the sense of Serre's theory.

Let $f(X)=\sum A'_wX^w$, $A'_w\in[q]$, be our polynomial of degree d defined over the finite field $[q]$, where $q=p^a$. It is obvious that $|w|\leqq d$, and $A'^{\,q}_w=A'_w$.

Let A_w be the Teichmuller representative of A'_w in K_a, so that $A_w^q=A_w$, and let

$$f_0(X)=\sum A_wX^w \tag{25}$$

be a polynomial over K_a such that its reduction modulus the local prime p in K_a gives $f(X)$. Let

$$F(X)=\prod_w\theta(A_wX^w) \tag{26}$$

$$F_0(X)=\prod_{j=0}^{a-1}{}^{\tau^j}F(X^{p^j}). \tag{27}$$

If $d=\deg(f)$, then (20) shows that

$$F(X)\in L(1/d), \tag{28}$$

$$F_0(X)\in L(p/qd). \tag{29}$$

Let α be the endomorphism of the Ω_0-linear space $L(p/qd)$ defined by

$$\alpha = \psi_p{}^a \circ F_0(X); \tag{30}$$

then the relation between the sum $S_m{}^*(f)$ and the theory we have outlined is expressed by the following equation (Dwork [1]):

$$(q^m - 1)^n \operatorname{Tr}(\alpha^m) = S_m{}^*(f), \tag{31}$$

which has of course a meaning because by (29) and (ii) the trace is well defined.

Now for any matrix α we have

$$\det(I - \alpha t) = \exp\left(-\sum_{m=1}^{\infty} \operatorname{Tr}(\alpha^m) t^m/m\right) \tag{32}$$

and introducing Dwork's δ operator defined by

$$\delta : g(t) \to g(t)^{\delta} = g(t)/g(qt) \tag{33}$$

we can write (31) as follows:

$$L^*(t,f)^{(-1)^{n-1}} = \det(I - \alpha t)^{\delta^n}. \tag{34}$$

All this, with some very small modifications, is the basis of Dwork's theory. Now we must explain that, in order to get information about $L^*(t,f)$ from equation (34), the most natural plan is to introduce differential operators in the space $L(1/d)$, which give rise to a sequence of finite dimensional spaces, whose dimension describes very accurately the structure of the function $L^*(t,f)$. However, this approach leads to difficult problems unless we assume some conditions of non-singularity of the polynomial f. As we are interested in properties valid for all functions $L^*(t,f)$, we shall follow a different and simpler method.

We have so far considered α as an Ω_0-linear endomorphism of $L(p/qd)$. On the other hand, $L(p/qd)$ is also an Ω_1-linear space, and we shall need to consider α as an Ω_1-linear endomorphism of $L(p/qd)$. We shall denote by $\det_{\Omega_0}(I - \alpha t)$ and $\det_{\Omega_1}(I - \alpha t)$ respectively, the characteristic functions of α accordingly to the first or the second point of view.

Let $c(m)$ denote the binomial coefficient $\binom{n+m-1}{m}$, and let

$$\det_{\Omega_0}(I - \alpha t) = \sum_{m=0}^{\infty} a_m t^m; \tag{35}$$

then we have the

LEMMA 2. *We have*

$$a_m \in \Omega_1 \text{ for all } m, \tag{36}$$

and the convex closure in the real (x,y) *plane of the points* $(m, \operatorname{ord}_q(a_m))$ *is contained in the convex closure of the points*

$$\Bigl(\sum_{j=0}^{m} c(j)\,;\,1/d\sum_{j=0}^{m} jc(j)\Bigr), \qquad m = 0, 1, \cdots. \tag{37}$$

Proof. We follow closely the method of Dwork [3], § 7 where a similar problem is considered.

First of all we have an easily established relation

$$\det_{\Omega_1}(I - \alpha t) = N_{\Omega_0/\Omega_1} \det_{\Omega_0}(I - \alpha t), \tag{38}$$

the norm from Ω_0 to Ω_1 being interpreted as the product of the conjugates of $\det_{\Omega_0}(I - \alpha t)$ over Ω_1, the automorphisms acting trivially on the variable t.

On the other hand, by equation (34) we have

$$\det_{\Omega_0}(I - \alpha t) = L^*(t,f)^{(-1)^{n-1}\delta^{-n}} \tag{39}$$

where δ^{-1} is the inverse operator of δ. By Lemma 1, the coefficients of the Taylor series for $L^*(t,f)$ are in the field obtained by adjoining the p-th roots of unity to the rational field, whence

$$\det_{\Omega_0}(I - \alpha t) \in \Omega_1\{t\}, \tag{40}$$

which proves (36) of Lemma 2. Also from equation (40) we have that the conjugates of $\det_{\Omega_0}(I - \alpha t)$ over Ω_1 are all equal, and we obtain

$$N_{\Omega_0/\Omega_1} \det_{\Omega_0}(I - \alpha t) = \det_{\Omega_0}(I - \alpha t)^a.$$

Combining this last equation with (38) we have

$$\det_{\Omega_0}(I - \alpha t)^a = \det_{\Omega_1}(I - \alpha t). \tag{41}$$

The Frobenius automorphism τ of the field Ω_0 acts trivially on Ω_1, whence we may look at τ as an Ω_1-linear endomorphism of $L(p/qd)$, again denoted by τ. Let α_0 be the Ω_1-linear endomorphism of $L(p/qd)$ defined by

$$\alpha_0 = \psi_p \circ \tau^{-1} \circ F(X)\,; \tag{42}$$

we shall prove that

$$\alpha = \alpha_0{}^a \tag{43}$$

where both members of (43) are considered as Ω_1-linear endomorphisms. In fact the Ω_1-linear endomorphism ${}^{\tau^h}F(X)$ of $L(p/qd)$ may be written as

$$ {}^{\tau^h}F(X) = \tau^h \circ F(X) \circ \tau^{-h}; \tag{44}$$

also, the Ω_1-linear endomorphisms ψ_p and τ^h commute, and it is easily seen that τ^a acts trivially on Ω_0.

On the other hand we have a general identity

$$\psi_p{}^a \circ \prod_{j=0}^{a-1} H_j(X^{p^j}) = \prod_{j=0}^{a-1} (\psi_p \circ H_{a-j-1}(X)), \tag{45}$$

and by definition we have

$$F_0(X) = \prod_{j=0}^{a-1} {}^{\tau^j}F(X^{p^j});$$

then equation (43) comes from (44), (45) and the two previous remarks.

It is obvious that

$$\det_{\Omega_1}(I - \alpha_0{}^a t^a) = \prod_{\zeta} \det_{\Omega_1}(I - \alpha_0(\zeta t))$$

where the product runs over all roots in Ω of $\zeta^a = 1$; hence from (41) and (43) we deduce

$$\det_{\Omega_0}(I - \alpha t^a)^a = \prod_{\zeta} \det_{\Omega_1}(I - \alpha_0(\zeta t)). \tag{46}$$

It follows without difficulty from equation (46) and the theory of Newton Polygons (see Dwork [2], Lemma 1.6 and [3], Lemma 7.1) that the convex closure of the points $(m, \mathrm{ord}_q(a_m))$ is obtained from the Newton Polygon of $\det_{\Omega_1}(I - \alpha_0 t)$ by reducing by a factor $1/a$ the ordinates and the abscissas of the latter. Hence in order to prove Lemma 2 we have to show that the Newton Polygon of $\det_{\Omega_1}(I - \alpha_0 t)$ is contained in the convex closure of the points

$$\Bigl(a\sum_{j=0}^{m} c(j); a/d \sum_{j=0}^{m} jc(j)\Bigr), \qquad m = 0, 1, 2, \cdots. \tag{47}$$

In order to do this we shall construct a matrix representation $(G_{u,i,v,j})$ over Ω_1 of the Ω_1-linear endomorphism α_0, indexed by pairs $((u,i),(v,j))$ where u and v are n-uples of non-negative integers, and where i, j run over the integers $1, 2, \cdots, a$. This matrix representation will have the following properties:

(A) $G_{u,i,v,j} = 0$ when the n-uple $pu - v$ contains some negative element;

(B) $\mathrm{ord}(G_{u,i,v,j}) \geqq 1/d \mid pu - v \mid /(p-1)$.

We consider Ω_0 as an Ω_1-module of dimension a over Ω_1, and we choose an integral basis $\xi_1, \xi_2, \cdots, \xi_a$ of Ω_0 over Ω_1, such that with respect to this basis the Ω_1-module Ω_0 is a p-adically direct sum, in the sense of Dwork [2], § 3(c), of a copies of Ω_1. By this we mean that the basis $\xi_1, \cdots, \xi_a$ has the property that for every $\beta_1, \cdots, \beta_a \in \Omega_1$

$$\operatorname{ord}(\beta_1\xi_1 + \cdots + \beta_a\xi_a) = \min_i \operatorname{ord}(\beta_i), \tag{48}$$

It is known that such a basis always exists.

Let $F(X) = \sum F_u X^u$, and let

$$F_u = \sum_{k=1}^{a} \phi_{u,k}\xi_k, \qquad \phi_{u,k} \in \Omega_1, \tag{49}$$

$$\tau^{-1}(\xi_j\xi_k) = \sum_{i=1}^{a} s_{ijk}\xi_i, \qquad s_{ijk} \in \Omega_1. \tag{50}$$

It is obvious that $\operatorname{ord}(\xi_i) = 0$ for every i, whence

$$\operatorname{ord}(\tau^{-1}(\xi_j\xi_k)) \geqq 0;$$

also by (28) we obtain

$$\operatorname{ord}(F_u) \geqq 1/d \mid u \mid/(p-1).$$

From these two inequalities and equation (48) we have

$$\operatorname{ord}(s_{ijk}) \geqq 0, \tag{51}$$

$$\operatorname{ord}(\phi_{u,k}) \geqq 1/d \mid u \mid/(p-1). \tag{52}$$

Let $\sum_u (\sum_{i=1}^{a} c_{u,i}\xi_i)X^u$, $c_{u,i} \in \Omega_1$, be a generic element of $L(p/qd)$; then the action of the Ω_1-linear endomorphism $\alpha_0 = \psi_p \circ \tau^{-1} \circ F(X)$ is described by the linear map

$$\alpha_0\colon\ c_{u,i} \to \sum_v \sum_{k=1}^{a} \phi_{pu-v,k} s_{ijk} c_{v,j}$$

and we obtain immediately a matrix representation $(G_{u,i,v,j})$ over Ω_1 of α_0, given by

$$G_{u,i,v,j} = \sum_{k=1}^{a} \phi_{pu-v,k} s_{ijk}. \tag{53}$$

Using (53) and the inequalities (51) and (52) it is immediate that this matrix representation satisfies conditions (A) and (B).

Let

$$\det_{\Omega_1}(I - \alpha_0 t) = \sum_{m=0}^{\infty} b_m t^m; \tag{54}$$

the Newton Polygon of this characteristic function is the convex closure of the points $(m, \operatorname{ord}(b_m))$.

On the other hand, by developing explicitly the determinant in (54) and writing $G(u, i, v, j) = G_{u,i,v,j}$ for simplicity, we have that b_m is a sum

$$b_m = \sum \pm \prod_{\nu=1}^{m} G(u_\nu, i_\nu, v_\nu, j_\nu)$$

where the summation is over all sets of m quadruplets $(u_\nu, i_\nu, v_\nu, j_\nu)$ such that the m pairs (u_ν, i_ν) are all *distinct,* and the m pairs (v_ν, j_ν) are a *permutation* of the m pairs (u_ν, i_ν); the signs in the sum are determined by m and the parity of the permutation. Using condition (B) we deduce without difficulty that

$$\operatorname{ord}(b_m) \geqq 1/d \min \sum_{\nu=1}^{m} | pu_\nu - v_\nu | / (p-1)$$

where the minimum is over all sets of m quadruplets $(u_\nu, i_\nu, v_\nu, j_\nu)$ satisfying the above conditions. Now we have

$$\sum_{\nu=1}^{m} | pu_\nu - v_\nu | \geqq \sum_{\nu=1}^{m} (p\,| u_\nu | - | v_\nu |) = (p-1) \sum_{\nu=1}^{m} | u_\nu |$$

because the m pairs (v_ν, j_ν) are a permutation of the m pairs (u_ν, i_ν). Thus we have proved that the Newton Polygon of $\det_{\Omega_1}(I - \alpha_0 t)$ is contained in the convex closure of the points

$$(m, 1/d \min \sum_{\nu=1}^{m} | u_\nu |) \tag{55}$$

where the minimum is now over all sets of m *distinct* pairs (u_ν, i_ν).

Then a simple computation (see Dwork [3], § 7, eq. 7.5 and the argument following it) shows that the convex closure of the points given by (55) coincides with the convex closure of the points

$$(a \sum_{j=0}^{m} c(j), a/d \sum_{j=0}^{m} jc(j)), \tag{47}$$

whence, by the remark preceding (47), Lemma 2 is proven.

COROLLARY. *$L^*(t, f)$ and hence $L(t, f)$ are rational functions of t.*

Proof. By the second part of Lemma 2 we have easily

$$\lim_{m \to \infty} 1/m \operatorname{ord}_q(a_m) = +\infty$$

whence $\det_{\Omega_0}(I - \alpha t)$ is a p-adically entire function. By equation (34) we deduce that $L^*(t, f)$ is a p-adically meromorphic function of t in all Ω. Now Lemma 1 tells us that the coefficients of the Taylor series of $L^*(t, f)$ are algebraic integers in K, and also that this Taylor series has a non zero radius of convergence with respect to every archimedean valuation of K. Thus it is clear how to apply Dwork's criterion of rationality (Dwork [1], Th. 3) and the corollary follows.

In view of this corollary we shall write

$$L^*(t, f)^{(-1)^{n-1}} = \prod_{h=1}^{r} (1 - \omega_h t) / \sum_{j=1}^{s} (1 - \eta_j t) \tag{56}$$

where ω_h^{-1} and η_j^{-1} are the zeros and poles of $L^*(t, f)^{(-1)^{n-1}}$.

IV°. Proof of Theorem 1. It is clear that (56) proves (i) of Theorem 1, except for the fact that the numbers ω_h, η_j are algebraic integers. This is an easy consequence of Lemma 1, and I am indebted to B. Dwork for this short proof (possibly not new): let k be the algebraic number field obtained by adjoining all the roots ω_h, η_j to the rational field Q, and suppose that ω_1, say, is not an integer in k. Let π be a prime ideal in k such that $\pi \mid (\omega_1^{-1})$, and let $| \ |_\pi$ be the non-archimedean valuation determined by π. Then Lemma 1 tells us that the Taylor series for $L^*(t, f)^{(-1)^{n-1}}$ has leading coefficient equal to 1, and the other coefficient are integers in k. Let k_π be the completion of k with respect to the prime π, and let Ω_π be the completion of the algebraic closure of k_π; then this Taylor series is convergent in Ω_π in the circle $|t|_\pi < 1$, and Krasner's theorem of the uniqueness of analytic continuation in Ω_π shows that this Taylor series converges in the circle $|t|_\pi < 1$ to the function $L^*(t, f)^{(-1)^{n-1}}$.

Now we get a contradiction because $\pi \mid (\omega_1^{-1})$ implies that $|\omega_1^{-1}|_\pi < 1$, and hence

$$|L^*(\omega_1^{-1}, f)^{(-1)^{n-1}}|_\pi = 1,$$

since the leading coefficient of its Taylor series is 1, and the other coefficients are integers in k; this is impossible because $L^*(\omega_1^{-1}, f)^{(-1)^{n-1}} = 0$. A similar argument applies to the η_j.

Now we prove statements (ii) and (iii) of Theorem 1.

Let $D_1(t)$ and $D_2(t)$ be the two entire functions in Ω defined by

$$D_1(t) = \prod_{h} \prod_{m=0}^{\infty} (1 - q^m \omega_h t)^{c(m)}, \tag{57}$$

$$D_2(t) = \prod_{j} \prod_{m=0}^{\infty} (1 - q^m \eta_j t)^{c(m)}; \tag{58}$$

we have

$$D_1(t)^{\delta^n} = \prod_h (1 - \omega_h t)$$

$$D_2(t)^{\delta^n} = \prod_j (1 - \eta_j t)$$

whence

(59) $$\det_{\Omega_0}(I - \alpha t) = D_1(t)/D_2(t),$$

as we can see from (34) and (56).

On the other hand in the proof of Lemma 2, Corollary, we have shown that $\det_{\Omega_0}(I - \alpha t)$ is an entire function in Ω, hence without poles. It follows that every zero of $D_2(t)$ is a zero of $D_1(t)$ with at least the same multiplicity. Let $\eta_1 = \eta_2 = \cdots = \eta_k$ be the set of all characteristic roots η_j which are equal to η_1, and consider the factor $(1 - q^m \eta_1 t)^{kc(m)}$ of $D_2(t)$. By our previous remark this must be a factor of $D_1(t)$, and letting $m \to \infty$ we see that we need at least k roots ω_h for expressing this factor as a product of factors of the type $(1 - q^m \omega_h t)^{c(m)}$. In other words, by reordering the roots ω_h we may write

(60) $$(q^m \eta_j)^{-1} = (q^{m_j} \omega_j)^{-1}, \qquad j = 1, 2, \cdots, k.$$

It is clear that $m \neq w_j$, whence the previous equation proves (ii) of Theorem 1.

The same argument proves (iii) of Theorem 1; in this case we remark that the root η_1, say, gives a factor $(1 - \eta_1 t)$ which must be again a factor of $D_1(t)$, whence there is a root ω_h such that

$$\eta_1 = q^{a_h} \omega_h, \qquad a_h > 0.$$

Without loss in generality we may take $h = 1$, because in any case the factor $(1 - q^{m+a_h} \omega_h t)^{c(m+a_h)}$ of $D_1(t)$ will cancel the factor $(1 - q^m \eta_1 t)^{c(m)}$ of $D_2(t)$ for every m, and (iii) of Theorem 1 is proven.

Now we prove (iv) of Theorem 1, which is in fact the most interesting part in Theorem 1. We need a simple lemma which relates the behaviour of the zeros of an entire function in Ω to the behaviour of the coefficients of its Taylor series. This lemma is the p-adic analogue of Jensen's formula for analytic functions of a complex variable; it is a fairly simple consequence of the theory of Newton Polygons.

Let

$$D(t) = 1 + \sum_{m=1}^{\infty} d_m t^m$$

be an entire function in Ω, and let $\{\rho_i\}$ be the sequence of the reciprocals of the zeros of $D(t)$, the ordering of this sequence being given by the valuation 'ord.' Then we have the

LEMMA 3. *With the previous notations we have for every real* x

$$\Sigma'(x - \mathrm{ord}(\rho_i)) = \max_{m \geqq 0} (xm - \mathrm{ord}(d_m)) \tag{61}$$

where Σ' *is a sum over all* i *such that* $\mathrm{ord}(\rho_i) \leqq x$.

Proof. The number $xm - \mathrm{ord}(d_m)$ is a maximum for some x if the point $(m, \mathrm{ord}(d_m))$ is a vertex of the convex closure of all such points, and without loss of generality we may also suppose that m has this property; for if not, then there exists m' with this property such that $xm' - \mathrm{ord}(d_{m'}) = xm - \mathrm{ord}(d_m)$. Thus in the proof of Lemma 3 we may suppose that $(m, \mathrm{ord}(d_m))$ is a vertex of the Newton Polygon of $D(t)$.

By Lemma 1.6 of Dwork [2] (in fact this lemma is easily extended to entire functions in Ω) we have that the vertices of the Newton Polygon of $D(t)$ are the points

$$(\Sigma' 1, \Sigma' \mathrm{ord}(\rho_i))$$

where Σ' is a sum over all i such that $\mathrm{ord}(\rho_i) \leqq y$ for some y. It follows immediately that

$$\max_m (xm - \mathrm{ord}(d_m)) = \max_y (\Sigma' x - \Sigma' \mathrm{ord}(\rho_i));$$

it is obvious that this maximum is attained for $y = x$ and Lemma 3 is proven.

Remark. Lemma 3 is certainly true with 'ord_q' instead of 'ord'; it is enough to divide both members by a and then put ax instead of x. Also it is enough to restrict m to integers m such that $(m, \mathrm{ord}(d_m))$ is a vertex of the convex closure of all such points.

COROLLARY. *The function* $\det_{\Omega_0}(I - \alpha t)$ *satisfies*

$$\sum_{h=1}^{r} \Sigma'(x - \mathrm{ord}_q(q^{m\omega_h}))c(m) - \sum_{j=1}^{s} \Sigma'(x - \mathrm{ord}_q(q^{m\eta_j}))c(m) \leqq d^{-1} \sum_{i \leqq dx} (dx - i)c(i), \tag{62}$$

where the sums Σ' *are over all* m *such that the summands are positive.*

Proof. Obvious from (59), Lemma 3, Lemma 2 and the second remark to Lemma 3.

Now we have

$$\sum_{m \leqq x} (x - m)c(m) = x^{n+1}/(n+1)! + (1/2)x^n/(n-1)! + O(x^{n-1}) \tag{63}$$

and it is clear that $\operatorname{ord}_q(q^m\omega_h) = m + \operatorname{ord}_q(\omega_h)$, whence

$$\Sigma'(x - \operatorname{ord}_q(q^m\omega_h))c(m) = x^{n+1}/(n+1)! + (n/2 - \operatorname{ord}_q(\omega_h))x^n/n! + O(x^{n-1})$$

and it follows that

$$\begin{aligned} &\sum_{h=1}^{r} \Sigma'(x - \operatorname{ord}_q(q^m\omega_h))c(m) - \sum_{j=1}^{s} \Sigma'(x - \operatorname{ord}_q(q^m\eta_j))c(m) \\ &\quad = (r-s)x^{n+1}/(n+1)! + (n(r-s)/2 - \sum_{h=1}^{r} \operatorname{ord}_q(\omega_h) \\ &\quad + \sum_{j=1}^{s} \operatorname{ord}_q(\eta_j))x^n/n! + O(x^{n-1}). \end{aligned} \tag{64}$$

On combining (64) with (62) and (63) we deduce

$$(r-s)x^{n+1}/(n+1)! + O(x^n) \leqq d^n x^{n+1}/(n+1)! + O(x^n)$$

and letting $x \to +\infty$ we obtain

$$(r-s) \leqq d^n,$$

that is (7) of Theorem 1.

It remains to prove the last statement of Theorem 1 concerning the case of equality in (7). For this we need

LEMMA 4. *Let $f(x)$ be a polynomial in one variable, defined over $[q]$, of degree d, $(d,p) = 1$. Then*

$$S_m(f) = -\sum_{h=1}^{d-1} \theta_h{}^m \tag{65}$$

where

$$|\theta_h| = \sqrt{q}. \tag{66}$$

Proof. By Weil's results (see for instance [11], last formula) we have

$$S_m(f) = -\sum_{h=1}^{D} \theta_h{}^m$$

where $|\theta_h| = \sqrt{q}$ and $D \leqq d-1$, provided only that

$$f(x) \neq h(x)^p - h(x)$$

for every polynomial $h(x)$ defined over the algebraic closure of $[q]$; this last condition is certainly satisfied if $(d,p) = 1$. We have to prove that $D = d-1$ when $(d,p) = 1$.

Let

$$L(t,f)=\sum_{m=0}^{D}A_m t^m;$$

we shall show that

$$A_{d-1}\neq 0,$$

whence $D=d-1$ because $D\leqq d-1$ and $L(t,f)$ is a polynomial of exact degree D. We give here a well known simple algebraic proof.

By equation (15), we have easily that

$$A_m=\sum e_p(\sigma\sum_{i=1}^{m}f(x_i))$$

the summation being over all positive 0-cycles $x_1+x_2+\cdots+x_m$ of degree m, defined over $[q]$. Hence

$$A_m=\sum_{g(\lambda)=0}e_p(\sigma\sum_{i=1}^{m}f(\lambda_i)) \tag{67}$$

where the summation is over all polynomials

$$g(x)=x^m+y_1x^{m-1}+\cdots+y_m \tag{68}$$

defined over $[q]$, and where the trace is from $[q]$ to $[p]$. Let us take $m=d-1$; the sum $\sum f(\lambda_i)$ is a symmetric function on the roots of $g(x)=0$, and it is very easy to see that if $f(x)=a_0x^d+\cdots+a_d$ then

$$\sum f(\lambda_i)=a_0\sum_{h=1}^{d-1}y_{d-h}(hy_h+P_h)+Q \tag{69}$$

where Q is a polynomial over $[q]$ in the variables $y_1,y_2,\cdots,y_{[d/2]-1}$, and where P_h is a polynomial over $[q]$ in the variables $y_1,\cdots,y_{h'}$ where

$$h'=\min(h,d-h)-1.$$

It follows that $\sum f(\lambda_i)$ is polynomial in $y_1,\cdots,y_{d-1}$ which is linear in $y_{d-1},y_{d-2},\cdots,y_{[d/2]+1}$ and quadratic in $y_{d/2}$ when d is even. On substituting (69) in (67) we see that the sum over y_{d-1} is 0 unless the coefficient of y_{d-1} vanishes. This coefficient by (69) is $da_0y_1+b_1$ for some constant b_1; as $(d,p)=1$ we have $da_0\neq 0$, whence the sum over y_{d-1} is o unless $y_1=-b_1/da_0$, in which case the sum is q times the sum (67) where now $y_1=-b_1/da_0$ and $y_{d-1}=0$.

Now we put $y_1=-b_1/da_0$, and we sum over y_{d-2}. This sum is again 0 unless $da_0y_2+P_2(-b_1/da_0)=0$, and so on. This process goes on up to $y_{d/2}$, where in the case d even we have a Gaussian sum, which is of modulus $\sqrt{q}$. The final result is

$$A_{d-1} = \epsilon(f)\, q^{(d-1)/2}$$

where $\epsilon(f)$ is a root of unity, and Lemma 4 is proven.

Remark. It is easy to see that if $(d,p) = p$ then $\deg L(t,f) < d-1$; in fact we have $\sigma x^p = \sigma x$ for every x, so that we may suppose that f has $(d,p) = 1$. For example taking $f(x) = x^p$ we have $S_m(x^p) = S_m(x)$ for every m; if $F(x) = f(x) + h(x)^p - h(x)$ where $h(x)$ is any polynomial defined over $[q]$, we must have

$$S_m(F) = S_m(f)$$

for all m.

COROLLARY. *When $n = 1$ and $(d,p) = 1$ we have $r = d$ and $s = 0$.*

Proof. $S_m^*(f) = S_m(f) - \epsilon^m$ where $\epsilon = e_p(\sigma f(0))$, and the trace σ is from $[q]$ to $[p]$.

Now we may complete the proof of Theorem 1. We take

$$f = x_1^d + \cdots + x_n^d;$$

then

$$S_m(f) = S_m(x^d)^n$$

and in view of Lemma 4, Corollary we have an example of a polynomial in n variables, of degree d with $(d,p) = 1$ and such that $r = d^n$, $s = 0$.

V°. Weil-type conjectures and *p*-adic estimates. Let us recall here some results about L-functions and some conjectures analogous to Weil's conjectures about Zeta functions. We consider for this moment only L-functions related to function fields in one variable and to non-trivial characters. The definitions will be given in the last section of this paper.

Let $\pi\colon Y \to X$ be a Galois covering of a curve X defined over $[q]$, with Galois group G, and let χ be a simple character of G. We assume that both Y and X are complete and normal. The L-function $L(t, Y, X, \chi)$ associated to this covering and to the character χ has the following properties: (Weil [12], pp. 72-83)

THEOREM A. *Let χ be a non-trivial simple character of G; then*

(i) *$L(t, Y, X, \chi)$ is a polynomial of degree*

$$D = (2g-2)\chi(1) + \deg \mathfrak{f}(\chi) \tag{70}$$

where g is the genus of X, $\chi(1)$ is the degree of an irreducible representation of the character, and where $\mathfrak{f}(\chi)$ is the Artin conductor of χ (it is a divisor on X). Also

(ii) *the zeros of $L(t,Y,X,\chi)$ have modulus $q^{-\frac{1}{2}}$;*

(iii) *we have a functional equation*

$$L(t,Y,X,\chi) = \epsilon(\chi) q^{D/2} t^D L(1/qt, Y, X, \bar{\chi}) \tag{71}$$

where $|\epsilon(\chi)| = 1$.

Of these results (i) is known as the Artin conjecture and (ii) as the Riemann hypothesis.

It is not very clear how to generalize this Theorem A to the case where $\dim X \geqq 2$; however we have some partial results which suggest the following hypothesis (Lang [5]). Let $L(t,Y,X,\chi)$ be an Artin L-function related to a Galois covering $\pi\colon Y \to X$, where Y and X are complete non-singular varieties of dimension n, defined over $[q]$, and let χ be a simple non-trivial character of the Galois group G of the covering π. Then we have the

HYPOTHESIS 1. *We have*

$$L(t,Y,X,\chi) = \prod_{h=1}^{2n-1} P_h(t)^{(-1)^{h-1}} \tag{72}$$

where the polynomials $P_h(t)$ have zeros of modulus $q^{-h/2}$, and satisfy

$$\deg P_h = \deg P_{2n-h}; \tag{73}$$

also we have a functional equation

$$L(t,Y,X,\chi) = \epsilon(\chi) q^{nD/2} t^D L(1/q^n t, Y, X, \bar{\chi}) \tag{74}$$

where $|\epsilon(\chi)| = 1$, and where D is the degree of $L(t,Y,X,\chi)$ as a rational function.

The statement about the zeros of $P_h(t)$ is the so-called Riemann hypothesis; equation (73) is analogous to the 'Poincaré duality' for the Betti numbers of a manifold. We do not know of any conjecture about D which generalizes (70) of Theorem A.

In the case of Artin L-functions $L(t,f)$ we have evidence for the truth of a similar hypothesis, which is much stronger than Hypothesis 1 but still rather vague.

HYPOTHESIS 2. *If f is a 'general' polynomial of degree d, in n variables, defined over $[q]$, and $(d,p) = 1$ we have:*

(i) $L(t,f)^{(-1)^{n-1}}$ *is a polynomial of exact degree* $(d-1)^n$, *whose zeros have modulus* $q^{-n/2}$;

(ii) $L(t,f)$ *satisfies a functional equation*

$$L(t,f)=\epsilon(f)q^{nD/2}t^{D}L(1/q^{n}t,-f) \tag{75}$$

where $\epsilon(f)$ *is a root of unity and* $D=(-1)^{n-1}(d-1)^n$.

We explain the meaning of 'general' by a concrete example. If $n=1$, then Lemma 4 shows that the hypothesis is true. Let $n=2$, $d=3$, so that $f(x_1,x_2)$ is a cubic polynomial. Let $F(x_1,x_2)$ be the homogeneous part of $f(x_1,x_2)$ of degree 3; then Davenport and I have proved that the hypothesis is true provided that the variety $F(X_1,X_2)=0$ in projective space (X_1,X_2) consists of three *distinct* points. It is obvious that this condition is a 'general' one.

We may even guess the following: let F be the homogeneous part of f, of degree d; then if the algebraic variety $F(X_1,\cdots,X_n)=0$ in projective space $(X_1,\cdots,X_n)$ is a regular hypersurface in the sense of Dwork [2], the hypothesis is verified. However the evidence for this result is very faint, and we cannot state it as a conjecture.

The fact that $L(t,f)^{(-1)^{n-1}}$ should be a polynomial of degree $(d-1)^n$ is suggested by the following consideration. By Theorem 1, we may expect that $r=d^n$, $s=0$ in most cases. Hence by (3) we may expect that the degree of $L(t,f)^{(-1)^{n-1}}$ as a rational function should be

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k}d^{n-k}=(d-1)^n.$$

We may relate to the conjectured functional equation (75) the following result:

THEOREM 2. *Suppose that* $S_m^*(f)$ *has* $r=d^n$, $s=0$. *Then we have*

$$\sum_{h=1}^{r}\operatorname{ord}_q(\omega_h)\geqq nd^{n-1}(d-1)/2 \tag{76}$$

and this is a best possible inequality.

Proof. We use (64) with $r=d^n$, $s=0$, and combine the resulting inequality with (63) and Lemma 3. We obtain

$$\begin{aligned}d^nx^{n+1}/(n+1)!+(nd^n/2-\sum_{h=1}^{r}\operatorname{ord}_q(\omega_h))x^n/n!+O(x^{n-1})\\ \leqq d^nx^{n+1}/(n+1)!+(nd^{n-1}/2)x^n/n!+O(x^{n-1})\end{aligned}$$

and letting $x\to+\infty$ (76) follows.

The example $f=x_1{}^d+\cdots+x_n{}^d$; $(d,p)=1$, shows that there may be equality in (76). We do not give the details of the proof of this fact, which depends on (4) and on the equation $A_{d-1}=\epsilon(f)q^{(d-1)/2}$ given in the course of the proof of Lemma 4.

We conclude this section with a theorem about p-adic estimates of the characteristic roots of the sum $S_m{}^*(f)$, and with a few remarks about the limitations of the method followed here.

THEOREM 3. *Let $S_m{}^*(f)$ be an exponential sum, where f is a polynomial in n variables of degree d, defined over $[q]$. Let ω_h, η_j be the characteristic roots of the sum $S_m{}^*(f)$, as defined in Theorem 1. Then $e_p(\sigma f(0,\cdots,0))$, where the trace is from $[q]$ to $[p]$, is one of the roots ω_h, and is the unique root of order 0. The other roots ω_h, η_j satisfy*

$$\operatorname{ord}_q(\omega_h)\geqq 1/d, \tag{77}$$

$$\operatorname{ord}_q(\eta_j)\geqq 1+1/d, \tag{78}$$

except for at most n roots η_j which must be all equal to $qe_p(\sigma f(0,\cdots,0))$.

Proof. Let $\det_{\Omega_0}(I-\alpha t)=\sum a_m t^m$; by Lemma 2 we have $\operatorname{ord}_q(a_0)=0$, $\operatorname{ord}_q(a_m)\geqq(m-1)/d$ for $m\geqq 1$. Also $\operatorname{ord}_q(a_1)=0$; in fact $a_1=\operatorname{Tr}(\alpha)$ whence

$$\operatorname{ord}_q(a_1)=\operatorname{ord}_q[(q-1)^n\operatorname{Tr}(\alpha)]=\operatorname{ord}_q(S_1{}^*(f)),$$

and $S_1{}^*(f)$ is a sum of $(q-1)^n$ p-th roots of unity. It follows that $S_1{}^*(f)$ is congruent to $(q-1)^n$, and hence to 1, modulus the maximal ideal in the ring of integers in Ω; this gives $\operatorname{ord}_q(S_1{}^*(f))=0$, and we have that $\operatorname{ord}_q(a_1)=0$.

Hence

$$\max_m(m/d-\operatorname{ord}_q(a_m))=1/d.$$

Now we use Lemma 3 with $x=1/d$. Clearly the sum over the roots η_j is empty, because by (iii) of Theorem 1 we must have $\operatorname{ord}_q(\eta_j)\geqq 1$. We obtain

$$\sum_{h=1}^{r}{}'(1/d-\operatorname{ord}_q(\omega_h))=1/d \tag{79}$$

and from the fact that $\operatorname{ord}_q(S_1{}^*(f))=0$ implies that one of the roots ω_h is of order 0, we deduce that there is exactly one root, say ω, or order 0, and we prove (77). Also (78) follows from (77) and (iii) of Theorem 1; the bound n for the exceptional roots η_j of order 1 comes from the fact that if there are k such roots then $(1-q\omega t)^k$ is a factor of the entire function $D_2(t)$ defined by (58), which must be also a factor of $D_1(t)$.

This proves Theorem 3 except for the fact that we have to prove that

$$\omega = e_p(\sigma f(0,\cdots,0)). \tag{80}$$

Let us write for brevity ρ for $e_p(\sigma f(0,\cdots,0))$, and assume that (80) is true for polynomials in fewer than n variables. We remark that by Lemma 4, Corollary, (80) is true for polynomials in one variable. Also if g is a polynomial in one variable by Theorem 3 we have

$$\operatorname{ord}_q(S_m(g)) \geqq m/d$$

because in this case, again by Lemma 4, the characteristic roots of $S_m(g)$ are the roots of $S_m{}^*(g)$ of positive order. (We have in Lemma 4 the condition $(d,p)=1$, but it is easy to see that these considerations apply also in the case $(d,p)=p$). It follows by induction on the number of variables of f that

$$\operatorname{ord}_q(S_m(f)) \geqq m/d \tag{81}$$

for every polynomial f.

By our induction hypothesis and (77) of Theorem 3 we deduce that if A is not the empty set then

$$\operatorname{ord}_q(S_m{}^*(f_A) - (-1)^{n-n(A)}\rho^m) \geqq m/d,$$

and also

$$\operatorname{ord}_q(S_m{}^*(f) - (-1)^n\omega^m) \geqq m/d,$$

so that by (3) we have easily

$$\operatorname{ord}_q(\omega^m - \rho^m) \geqq \min(m/d, \operatorname{ord}_q(S_m(f))) = m/d;$$

in particular

$$\operatorname{ord}_q(\omega - \rho) > 0.$$

On the other hand we have

$$\omega^m - \rho^m = (\omega-\rho)[m\rho^{m-1} + \binom{m}{2}\rho^{m-2}(\omega-\rho) + \cdots]$$

and from the fact that $\operatorname{ord}_q(\omega-\rho) > 0$ we deduce that

$$\operatorname{ord}_q(\omega^m - \rho^m) = \operatorname{ord}_q(\omega-\rho) \text{ if } (m,p)=1.$$

Letting $m \to \infty$ we obtain

$$\operatorname{ord}_q(\omega-\rho) = +\infty$$

i. e. $\omega = \rho$, and equation (80) is proven.

It may be of interest to examine the limitations of our method. Let $H(t)=\sum_{m=0}^{\infty} b_m t^m$ be p-adically entire, and suppose that $H(t)^{\delta^n}$ is a rational function of t. Then we have the following result:

THEOREM 4. *The degree of* $H(t)^{\delta^n}$ *as a rational function of* t *is given by*

$$\deg H(t)^{\delta^n}=k^{-n} \tag{82}$$

where k *is defined by*

$$c_n k=\liminf m^{-(1+1/n)} \operatorname{ord}_q(b_m) \tag{83}$$

and

$$c_n=(n/n+1)(n!)^{1/n}.$$

Proof. Let $H(t)^{\delta^n}=\prod_{h=1}^{r}(1-\lambda_h t)/\prod_{j=1}^{s}(1-\rho_j t)$; then

$$H(t)=\Delta_1(t)/\Delta_2(t)$$

where

$$\Delta_1(t)=\prod_{h=1}^{r}\prod_{m=0}^{\infty}(1-q^m\lambda_h t)^{c(m)}$$

$$\Delta_2(t)=\prod_{j=1}^{s}\prod_{m=0}^{\infty}(1-q^m\rho_j t)^{c(m)}.$$

On applying Lemma 3 to the entire function $H(t)$ we deduce that

$$(r-s)x^{n+1}/(n+1)!\sim\max_m(xm-\operatorname{ord}_q(b_m));$$

by (83) we have for a suitable sequence $x'\to\infty$ that

$$\max_m(x'm-\operatorname{ord}_q(b_m))\sim k^{-n}x'^{n+1}/(n+1)!$$

and Theorem 4 follows immediately.

Now we can see the main limitation of our method. Let $P(t)$ be a polynomial of degree d, and let

$$H_0(t)=H(t)P(t).$$

Then $H_0(t)$ is entire and satisfies the conditions of Theorem 4; let r_0 and s_0 respectively be the degree of the numerator and the denominator of $H_0(t)^{\delta^n}$. We have

$$r_0-s_0=r-s;$$

however we may have

$$r_0+s_0=r+s+d2^n.$$

Thus, in order to have more refined results from our method, we need some further condition about the entire function $H(t)$.

VI°. Artin-Schreier coverings of curves. In this section we prove some results about exponential sum 'along a curve,' which generalize the sums $S_m(f)$. We shall prove first a theorem in the complete non-singular case, and then show how to handle the general case where the given curve may be singular or even reducible. We shall follow closely (except for the language) Weil's method [11], where he considers exponential sums 'along a line.' Our generalization is straightforward but we believe that the results will be general enough for application to number theory, where we need results valid also in the singular case.

Let X be a complete non-singular curve of genus g, defined over the finite field $k=[q]$, let $k(x)$ be its function field, and let $\bar{k}(x)$ be the function field of X considered as a curve defined over the algebraic closure $\bar{k}$ of $[q]$. Let $R(x)\in k(x)$ be a rational function on X, satisfying the condition

$$R(x)\neq h^p-h \text{ for } h\in\bar{k}(x). \tag{84}$$

We may identify $R(x)$ with a rational map

$$R: X\rightarrow A^1$$

of X into the affine line A^1 defined over $[q]$; thus we may speak about the value of the function $R(x)$ at a point x of the curve X.

Let X_m be the set of points of X defined over $[q^m]$, and let

$$S_m(R,X)=\sum_{x\in X_m}{}' e_p(\sigma R(x)), \tag{85}$$

where the accent in the summation means that the poles of $R(x)$ are excluded in it. This exponential sum is related to an Artin-Schreier covering of X, which we are going to define.

Let X' be the curve defined by the equation

$$X': z^p-z=R(x); \tag{86}$$

then we have a Galois covering $\pi: X'\rightarrow X$, with Galois group Z/pZ acting on X' by means of the map $g: z\rightarrow z+g$. Let Y be a normalization of X'; then we have an everywhere defined birational map

$$h: Y\rightarrow X',$$

and from this we obtain the covering

$$f: Y\rightarrow X, \tag{87}$$

which is the Artin-Schreier covering of X related to the function $R(x)$. We remark that both X and Y are non-singular complete curves, defined over $[q]$.

The L-function related to this covering is, as we shall show in the last section of this paper, given by

$$L(t,R,X)=\exp\left(\sum_{m=1}^{\infty}S_m(R,X)t^m/m\right) \tag{88}$$

for $|t|<q^{-1}$.

Now let $(R)_\infty$ be the divisor of poles of $R(x)$ on X, and write

$$(R)_\infty=\sum_{i=1}^{t}d_ix_i, \tag{89}$$

where the x_i are points on X, and d_i is the multiplicity of the pole of $R(x)$ at x_i. Then we have the following

THEOREM 5. *We have*

$$S_m(R,X)=-\sum_{h=1}^{D}\theta_h{}^m, \tag{90}$$

where

$$|\theta_h|=\sqrt{q} \tag{91}$$

and

$$D\leqq 2g-2+t+\deg(R)_\infty. \tag{92}$$

Moreover, we have equality in (92) *if and only if* $(d_i,p)=1$ *for all* $i=1,\cdots,t$.

Proof. By (i) of Theorem A (the Artin conjecture), $L(t,R,X)$ is a polynomial in t, which implies the decomposition (90). Equation (91) is the Riemann hypothesis for $L(t,R,X)$ i.e., (ii) of Theorem A. We have to prove (92) and the statement concerning the case of equality in it. This amounts to finding the degree of the conductor of the character $e_p(u)$ of the Galois group of the previous covering.[1]

Let K be the algebraic number field of p-th roots of unity; the function $L(t,R,X)$ is a polynomial over K of degree D, and its conjugates over the rational field Q are the functions

$$\exp\left(\sum_{m=1}^{\infty}S_m(hR,X)t^m/m\right)$$

for $h=1,2,\cdots,p-1$.

---

[1] The degree of the conductor has been determined by Hasse [4]; see also Serre [8], Ch. VI, § 12, ex. 2. We give here another proof, based on a reduction to Lemma 4.

On the other hand we have (orthogonality of characters)

$$\sum_{h=0}^{p-1} e_p(h\sigma x) = \begin{cases} p \text{ if } x = 0 \\ 0 \text{ otherwise} \end{cases}$$

whence we obtain

$$N_{K/Q} L(t, R, X) = Z(t, Y)/Z(t, X), \tag{93}$$

where the norm is from K to Q, the automorphisms acting trivially on the variable t, and where $Z(t, V)$ denotes the Weil Zeta function of a variety V.

By a well known result of Weil [12], we have that if C is a completely non-singular curve of genus γ then the degree of the rational function $Z(t, C)$ is

$$\deg Z(t, C) = 2\gamma - 2,$$

whence if g' and g denote the genus of Y and X, we have from (93) the equation.

$$(p-1)D = (2g'-2) - (2g-2). \tag{94}$$

In order to find g', we use the Hurwitz-Zeuthen formula modified for application to the case of a field with non-zero characteristic. We follow here Serre [9], Ch. VI.

Every point $y \in Y$ has a local ring which is a discrete valuation ring (y is a simple point of Y), of field of fractions $k(y)$, and residue class field k; let v_y be the corresponding valuation, and let L_y denote the completion of $k(y)$ for the valuation v_y. In the same way if $x \in X$ we may define the valuation v_x and the complete field K_x which is the completion of $k(x)$ for v_x.

Now consider the covering $f\colon Y \to X$ of X. Its Galois group G is cyclic of order p, whence if $y \in Y$ the decomposition group of y is either G or 0 (the unit of $G = Z/pZ$). Thus we see that the decomposition group of y is G if and only if y lies above a point $x \in X$ which ramifies completely in the covering f. Now the covering f can ramify only at the poles x_i of $R(x)$, and if $(d_i, p) = 1$ there is ramification, and the ramification index is p, which is equal to the characteristic of $k(x)$.

Thus in order to apply the Hurwitz-Zeuthen formula we need to consider the ramification groups at the various points of y. The extension L_y/K_x is a Galois extension, with Galois group the decomposition group of y; let $G_h(y)$ be the h-th ramification group of the decomposition group of y in this extension, and let $\operatorname{Card} G_h(y)$ be the order of $G_h(y)$. Then we have

$$(2g'-2) = p(2g-2) + \sum_{y \in Y} \sum_{h=0}^{\infty} (\operatorname{Card} G_h(y) - 1). \tag{95}$$

Without loss of generality we may consider in (95) only points y lying above poles of $R(x)$. Let x be a pole of $R(x)$, of multiplicity d, and suppose that x ramifies in the covering f, and let y be the corresponding point on Y. As G is cyclic of order p, and as $G_h(y)$ must be a subgroup of G, we deduce that $G_h(y) = G$ or 0, whence by definition of ramification group we have that $G_h(y) = G$ if and only if

$$h \leqq v_y(s(u) - u), \qquad s \neq 0, \tag{96}$$

where u is a local uniformizing parameter at y, and where $s(u)$ denotes the action of the element $s \in G$ on u. Hence if y_i is the point of Y lying above x_i, we have from (95) and (96) that

$$(2g' - 2) \leqq p(2g-2) + \sum_{i=1}^{t} (v_{y_i}(s(u) - u) + 1)(p-1), \tag{97}$$

and equality holds in (97) if and only if the covering f is ramified at every pole of $R(x)$. This is indeed the case when $(d_i, p) = 1$ for all i.

Thus in order to prove Theorem 5, in view of (94) and (95) we have only to prove that if $s \neq 0$, $s \subset G$, then

$$v_y(s(u) - u) \leqq d, \tag{98}$$

with equality if and only if $(d, p) = 1$.

The question here is purely local; the extension L_y of K_x is given by means of an Artin-Schreier equation $z^p - z = a$, where $a \in K_x$ is an element of valuation $v_x(a) = -d$. Taking a local uniformizing parameter T^{-1} at the point x, we have that the complete field K_x is isomorphic to the field $[q]((T^{-1}))$ of formal Laurent series of T^{-1}, with coefficients in $[q]$, and with a finite number of negative powers of T^{-1}. Also, if $(d, p) = 1$, we may choose the local uniformizing parameter T^{-1} such that $a = T^d$, while if $(d, p) = p$ we may choose it such that $a = P(T)$, where $P(T)$ is a polynomial of degree d. However, in this new situation, we are in the case where X is the projective line defined over $[q]$, and the rational function $R(x)$ is a polynomial in T of degree d. In this case we have $g = 0$, $t = 1$, $x_1 = \infty$, whence by (94) and (97) we obtain

$$(p-1)D' = 2g'',$$
$$2g'' - 2 = -2p + (p-1) + v_{y'}(s(u) - u)(p-1)$$

where g'' is the genus of the curve Γ, $z^p - z = T^d$ if $(d, p) = 1$, and $z^p - z = P(T)$ if $(d, p) = p$, where D' is the degree of the L-function associated with this covering Γ of the projective line, and where y' is the point on the normalization of Γ lying above the pole $x_1 = \infty$ of T^d or $P(T)$.

By the local character of our construction we have

$$v_y(s(u) - u) = v_{y'}(s(u) - u);$$

on the other hand by Lemma 4 and the Remark to Lemma 4 we have $D' = d - 1$ if $(d, p) = 1$, and $D' < d - 1$ if $(d, p) = p$. It follows immediately that

$$v_{y'}(s(u) - u) \leqq d$$

with equality if and only if $(d, p) = 1$. This proves (98) and, with it, Theorem 5.

Now we give a general inequality for exponential sums $S_m(R, X)$ where X is a projective curve which may be singular and reducible.

Let X be a projective curve of degree d_1 defined over the finite field $k = [q]$, embedded in n-dimensional projective space P^n over $[q]$. Let $R(X_1, \cdots, X_{n+1})$ be a homogeneous rational function in P^n, defined over $[q]$, and let d_2 be the degree of its numerator. Then we define

$$S_m(R, X) = \sum_{x \in X_m}' e_p(\sigma R(x))$$

in exactly the same way as (85).

Let $\Gamma_1, \Gamma_2, \cdots, \Gamma_s$ be the absolutely irreducible components of X, and suppose that the following condition is verified:

(A) for every homogeneous rational function $h = h(X_1, \cdots, X_{n+1})$ defined over the algebraic closure $\bar{k}$ of $k = [q]$, the function $R - h^p + h$ does not vanish identically on any component Γ of X.

Then we have the following

THEOREM 6. *If R satisfies conditions* (A) *then*

$$|S_m(R, X)| \leqq (d_1^2 + 2d_1 d_2 - 3d_1)(\sqrt{q})^m + d_1^2. \tag{99}$$

In particular this is true if $d_1 d_2 < p$ and R is not a constant on any component Γ of X.

Proof. Suppose first that X is absolutely irreducible. Let X^* be a normalization of X, so that we have an everywhere defined birational map

$$h \colon X^* \to X;$$

let $R(x)$ be the rational function on X induced by R (we may suppose that $R(x)$ is defined almost everywhere, otherwise the sum $S_m(R, X)$ is empty), which we identify with a rational map, again denoted by R

$$R \colon X \to A^1$$

of X into the affine line defined over $[q]$. The composite map $R \circ h$ gives rise to a rational function $r(x)$ on X^*, and we have

$$r(x) = R(h(x)) \quad \text{for } x \in X^*.$$

Now we have

$$(100) \qquad |S_m(R, X) - S_m(r, X^*)| \leqq d_1^2.$$

In fact the inverse map h^{-1} is defined except at a finite number of points of X, and the total number of points of X^* such that $h^{-1}(h(x))$ is not defined is at most $(d_1 - 1)(d_1 - 2) - 2g$, where g is the genus of X^*.

Now we are in position of applying Theorem 5 to the sum $S_m(r, X^*)$, because by condition (A) we see that $r(x)$ satisfies condition (84). Hence

$$(101) \qquad S_m(r, X^*) \leqq (d_1^2 - 3d_1 + 2\deg(r)_\infty)(\sqrt{q})^m$$

because $2g - 2 \leqq d_1^2 - 3d_1$ and obviously $t \leqq \deg(r)_\infty$. By Bézout theorem we have $\deg(r)_\infty \leqq d_1 d_2$, whence by (100) and (101) we obtain Theorem 6 in the case where X is absolutely irreducible.

Let X be irreducible over $[q]$, but not absolutely irreducible, so that

$$X = \Gamma_1 + \Gamma_2 + \cdots + \Gamma_s$$

is the sum of an absolutely irreducible curve Γ_1 of degree d_1/s, defined over $[q^s]$, and of its conjugates over $[q]$. Then we have

$$|S_m(R, X) - \sum_{h=1}^{s} S_m(R, \Gamma_h)| \leqq (1 - s^{-1}) d_1^2$$

because the total number of points of X common to two or more components Γ_h is at most $(1 - s^{-1}) d_1^2$.

It is well known that if $s \nmid m$ then there are at most $(d_1/s)^2$ points of Γ_h defined over $[q^m]$; for if $x \in \Gamma_h$ and x is defined over $[q^m]$ then x belongs to the conjugates of Γ_h over $[q^m]$. Now the smallest field of definition for Γ_h containing $[q]$ is $[q^s]$, whence if $s \nmid m$ then there is at least one conjugate of Γ_h over $[q^m]$, different from Γ_h. Hence

$$|S_m(R, X)| \leqq d_1^2/s \quad \text{if } s \nmid m,$$

while if $s \mid m$ by Theorem 6 applied to the sums $S_m(R, \Gamma_h)$ we obtain

$$|S_m(R, X)| \leqq (d_1^2/s - 3d_1 + 2d_1 d_2)(\sqrt{q})^m + d_1^2.$$

This proves Theorem 6 in the case where X is irreducible over $[q]$.

Now suppose that

$$X = a_1 C_1 + \cdots + a_u C_u$$

is the decomposition of X in distinct curves C_h defined and irreducible over $[q]$, of degree δ_h and multiplicity a_h. Then the same argument as before gives

$$| S_m(R, X) - \sum_{h=1}^{u} S_m(R, C_h) | \leqq \sum_{h \neq k} \delta_h \delta_k,$$

and by Theorem 6 applied to the curves C_h we have

$$| S_m(R, X) | \leqq (\sum \delta_h^2 + (2d_2 - 3) \sum \delta_h) (\sqrt{q})^m + (\sum \delta_h)^2;$$

obviously $\sum \delta_h \leqq d_1$ and we obtain Theorem 6 in the general case.

Remark. It is possible to show that the constant $d_1^2 - 3d_1 + 2d_1d_2$ cannot be improved (use Theorem 5 and the argument of Lang [6], pp. 138-139); the constant term d_1^2 in (99) may be improved to $(d_1 - 1)(d_1 - 2)/2$.

VII°. Exponential sums and Artin *L*-functions. We give here a short explanation of the connection between the theory of exponential sums in many variables and the general theory of Artin L-series. The main result of this section, which we believe to be new, is that the Riemann hypothesis for Weil Zeta functions implies the truth of the Riemann hypothesis for Artin L-functions related to coverings of the type Artin-Schreier. This is straightforward in the case of Artin L-functions of curves, once we have the Artin conjecture; however in the general case these L-functions are only rational functions, and this deduction is not immediate.

Let X be an absolutely irreducible normal variety defined over $[q]$, and let K be its function field. Let L be a Galois extension of K of degree n, with Galois group G. Let Y be a normalization of X in L, defined over $[q]$; then Y is a covering of degree n of X, and the group G defines a group of automorphisms of Y, again denoted by G, such that X can be identified with the quotient variety Y/G.

This gives a projection

$$\pi : Y \rightarrow X,$$

such that in general a point x of X is the image under π of n points of Y; we say that x is unramified in Y if x is the image of n distinct points of Y.

Let x be a point of X, and let $[q^m]$ be its smallest field of definition containing $[q]$. Let $\mathfrak{p} = x + x^q + \cdots + x^{q^{m-1}}$ be the prime 0-cycle of degree m of X associated with x, and suppose that $\mathfrak{p}$ is unramified in Y, i.e. x is unramified. Then $\mathfrak{p}$ is the image of r prime 0-cycles $\mathfrak{P}_1, \cdots, \mathfrak{P}_r$ of Y, defined over $[q]$. Let $\phi : x \rightarrow x^q$ be the Frobenius morphism of X; since x is defined over $[q^m]$, we have $\phi^m(x) = x$, hence in an obvious way we may write $\phi^m(\mathfrak{p}) = \mathfrak{p}$, meaning with the sign of equality that ϕ^m acts trivially on $\mathfrak{p}$. For the prime 0-cycles $\mathfrak{P}_i$ we have $\phi^m(\mathfrak{P}_i) \to \mathfrak{P}_i$, that is ϕ^m acts on $\mathfrak{P}_i$ by permuting its points. Now as $\mathfrak{p}$ is left fixed by ϕ^m, there exists $g_i \in G$ such that $\phi^m(\mathfrak{P}_i) = g_i(\mathfrak{P}_i)$. This element g_i is called the Frobenius substitution of $\mathfrak{P}_i$ in Y, and it is denoted by the symbol $(\mathfrak{P}_i, L/K) = g_i$. Also it is possible to show that the substitutions g_i all belong to a same conjugacy class of G.

Now let ρ denote an irreducible matrix representation of G in the general linear group $GL(M)$ where M is a finite dimensional vector space over the complex field; then to every $g \in G$ there corresponds a matrix $\rho(g)$, and the trace of this matrix is a character χ of G:

$$\chi(g) = \mathrm{Tr}(\rho(g)).$$

It is well known that the characteristic equation $\det(I - \rho(g)t)$ is invariant under inner automorphisms of G, whence from the remark that the remark that the substitutions g_i all belong to a same conjugacy class of G we have that the characteristic function $\det(I - \rho((\mathfrak{P}_i, L/K))t)$ is independent of i; we shall write $\det(I - \rho_{\mathfrak{p}}t)$ for this determinant.

Then the Artin L-function related to the Galois covering $\pi\colon Y \to X$ and to the character χ of the Galois group G of π, is given by

$$L(t, Y, X, \chi) = \prod_{\mathfrak{p}\text{ unram.}} \det(I - \rho_{\mathfrak{p}} t^{\deg(\mathfrak{p})})^{-1} \tag{102}$$

where the product is over all unramified prime 0-cycles of X, defined over $[q]$.

Now we show how to apply these considerations in order to get our function $L(t, f)$. We take for X the affine n-dimensional space A^n of points $x = (x_1, \cdots, x_n)$, and defined over $[q]$, and the covering $\pi\colon Y \to X$ is given by

$$Y\colon z^p - z = f(x_1, \cdots, x_n), \tag{103}$$

the projection π being given by $\pi(z, x) = x$. Then if f is such that $f \neq h^p - h$ for every polynomial h defined over the algebraic closure of $[q]$, the covering $\pi\colon Y \to X$ is a Galois covering of degree p, with Galois group $G = Z/pZ$.

Let $(z, x) \in Y$; then if $g \in G$ we have $g(z, x) = (z + g, x)$ because $(z + g)^p - (z + g) = z^p - z$ for every g. If $\mathfrak{p} = x + x^q + \cdots + x^{q^{m-1}}$ is a prime 0-cycle of degree m of A^n, defined over $[q]$, we have

$$\phi^m(z, x) = (z^{q^m}, x^{q^m}) = (z^{q^m}, x);$$

also $\mathfrak{p}$ is unramified in Y and it follows that the Frobenius substitution is given by the equation

$$(z + g, x) = (z^{q^m}, x),$$

that is $g = z^{q^m} - z$. On the other hand we have

$$z^{q^m} - z = (z^p - z) + (z^p - z)^p + \cdots + (z^p - z)^{q^m/p}$$
$$= f + f^p + \cdots + f^{q^m/p} = \sigma f,$$

whence $g = \sigma f$. The character χ in this case is given by $\chi(g) = e_p(g)$, and it follows immediately that the Artin L-function is given by

$$\prod_{\mathfrak{p}} (1 - e_p(\sigma f(\mathfrak{p})) t^{\deg(\mathfrak{p})})^{-1}$$

that is the function $L(t,f)$.

More generally, we may consider the Artin L-function

$$(104) \qquad L(t,R,X) = \exp(\sum_{m=1}^{\infty} S_m(R,X) t^m/m),$$

where R is a rational function on X defined over $[q]$ and satisfying condition (84), and where $S_m(R,X)$ is defined as in (85). This L-function is associated with the covering $\pi\colon Y \to X$ where

$$(105) \qquad Y\colon z^p - z = R(x), \quad x \in X.$$

Let $L(t,Y,X,\chi)$ be an Artin L-function related to a Galois covering $\pi\colon Y \to X$ defined over $[q]$. If $[q^a]$ is a finite extension of the field $[q]$, then the covering $\pi\colon Y \to X$ is defined also over $[q^a]$, and it remains a Galois covering. Thus we obtain a new L-function, and we shall write $L(t,Y,X,\chi,q^a)$ when necessary in order to put in evidence that the L-function depends on the field of definition of the covering π. The relationship between L-functions corresponding to the same covering but with different fields of definition is well known and quite simple:

$$(106) \qquad L(t^a,Y,X,\chi,q^a) = \prod_{\zeta} L(\zeta t,Y,X,\chi,q),$$

where the product is over all roots of $\zeta^a = 1$.

Now we return to the functions $L(t,R,X)$. Let $\lambda \in [q^a]$, and let $Y(\lambda)$ denote the covering of X

$$(107) \qquad Y(\lambda)\colon z^p - z = R(x) + \lambda, \quad x \in X,$$

where $[q^a]$ is the field of definition of $Y(\lambda)$, X and the covering $\pi\colon Y(\lambda) \to X$. We suppose that this covering is unramified, and we can do this by taking for X a non-complete variety, i. e. X — (the ramified points).

THEOREM 7. (i) *The function $L(t,R,X)$ is a rational function of t. Also there exists a suitable power p^μ of the degree p of the covering $\pi\colon Y \to X$*

such that (ii) *if* θ *is any characteristic root of* $L(t,R,X)$ *then there is* $\nu \leqq \mu-1$ *and* $\lambda \in [q^{p^\nu}]$ *such that*

$$\theta^{p^\mu} = \omega^{p^{\mu-\nu}}, \tag{108}$$

where ω *is a characteristic root of the Weil Zeta function* $Z(t,Y(\lambda),q^{p^\nu})$. *Also we have*

$$Z(t,Y(\lambda),q^{p^{\nu+1}}) = Z(t,Y,q^{p^{\nu+1}}). \tag{109}$$

COROLLARY. *Suppose that the following statements are true for Weil Zeta-functions.*

(i) *each characteristic root has modulus which is a non-negative integral power of* $\sqrt{q}$;

(ii) *the modulus of each root is the same as that of each of its conjugacy over* Q.

Then statements (i) *and* (ii) *remain true for Artin L-functions corresponding to Artin-Schreier coverings.*

Proof. (i) of Theorem 7 has been proved by Serre (see his report [10]), and in fact it is a consequence of Lemma 2, Corollary. Without loss of generality we may suppose X to be affine; then let $h_i(x_1,\cdots,x_n)=0$, $i=1,\cdots,s$ be a system of defining equations for X, and let

$$R=f(x_1,\cdots,x_n)/g(x_1,\cdots,x_n),$$

where f and g are polynomials defined over $[q]$. Let F be the polynomial over $[q]$ in the variables $x_1,\cdots,x_n,t_1,\cdots,t_s,z_1,z_2$ defined by

$$F=t_1h_1+\cdots+t_sh_s+z_1f+z_2-z_1z_2g. \tag{110}$$

Then we have

$$q^{(s+1)m}S_m(R,X)=S_m(F) \tag{111}$$

as we can see on summing in $S_m(F)$ over the variables t_i and z_j. It follows that

$$L(q^{s+1}t,R,X)=L(t,F), \tag{112}$$

and in view of Lemma 2, Corollary, we have (i) of Theorem 7.

Let $\lambda \in [q]$, and let $\rho=e_p(\sigma\lambda)$; it is obvious that

$$S_m(R+\lambda,X)=\rho^m S_m(R,X)$$

whence

$$L(t,R+\lambda,X)=L(\rho t,R,X). \tag{113}$$

Let K denote the algebraic number field of p-th roots of unity; then all these L-functions are elements of $K(t)$, and we have for every $u \in [p]$:

$$\begin{aligned} N_{K/Q}L(t, R+u\lambda, X) &= N_{K/Q}L(\rho^u t, R, X) \\ &= \prod_{v=1}^{p-1} L(\rho^{uv}t, vR, X) = Z(t, Y(u\lambda))/Z(t, X), \end{aligned} \tag{114}$$

as we can see using (113) and equation (93), which is true also in the case where X is a variety of any dimension. It follows that if $\sigma\lambda \neq 0$, so that ρ is a primitive p-th root of unity, then

$$\begin{aligned} &\prod_{u=0}^{p-1} Z(\rho^{-u}t, Y(u\lambda))/Z(\rho^u t, Y(\lambda)) \\ &\qquad = \prod_{u=0}^{p-1} \prod_{v=1}^{p-1} L(\rho^{u(v-1)}t, vR, X)/L(\rho^{u+v}t, vR, X). \end{aligned}$$

For a fixed $v \neq 1$ the product over u is equal to 1, and we obtain

$$\prod_{u=0}^{p-1} L(t, R, X)/L(\rho^u t, R, X) = \prod_{u=0}^{p-1} Z(\rho^{-u}t, Y(u\lambda))/Z(\rho^u t, Y(\lambda))$$

whence by (106) we have the important identity

$$\begin{aligned} &L(t, R, X, q)^p/L(t^p, R, X, q^p) \\ &\qquad = \{\prod_{u=0}^{p-1} Z(\rho^{-u}t, Y(u\lambda), q)\}/Z(t^p, Y(\lambda), q^p). \end{aligned} \tag{115}$$

Now we prove (109). Let $\lambda \in [q^{p^\nu}]$, and consider $Z(t, Y(\lambda), q^{p^{\nu+1}})$. Take $\lambda' \in [q^{p^{\nu+1}}]$ such that $\lambda'^p - \lambda' = \lambda$; then the number of solutions over $[q^{mp^{\nu+1}}]$ of the equation

$$z^p - z = R(x) + \lambda, \quad x \in X,$$

is the same as the number of solutions of

$$(z - \lambda')^p - (z - \lambda') = R(x), \quad x \in X.$$

The map $z \to z - \lambda'$ is a translation in the additive group of $[q^{mp^{\nu+1}}]$ for every m, whence the number of points of $Y(\lambda)$ defined over $[q^{mp^{\nu+1}}]$ is the same as the number of points of Y defined over the same field. By definition of Weil Zeta functions, (109) follows immediately.

Thus (115) becomes

$$\begin{aligned} &L(t, R, X, q)^p/L(t^p, R, X, q^p) \\ &\qquad = \{\prod_{u=0}^{p-1} Z(\rho^{-u}t, Y(u\lambda), q)\}/Z(t^p, Y, q^p), \end{aligned} \tag{116}$$

where $\rho = e_p(\sigma\lambda)$ and $\lambda \in [q]$ is such that $\sigma\lambda \neq 0$.

Now we apply (116) to the L-function $L(t,R,X,q^p)$, and so on. We obtain the identity

$$\begin{aligned} &L(t,R,X,q)^{p^\mu}/L(t^{p^\mu},R,X,q^{p^\mu}) \\ (117) \qquad &=\prod_{\nu=0}^{\mu-1}\{\prod_{u=0}^{p-1} Z(\rho_\nu{}^{-u}t^{p^\nu},Y(u\lambda_\nu),q^{p^\nu})\}^{p^{\mu-1-\nu}}/Z(t^{p^{\nu+1}},\ Y,q^{p^{\nu+1}})^{p^{\mu-1-\nu}}, \end{aligned}$$

where λ_ν is any element of $[q^{p^\nu}]$ such that

$$(118) \qquad \sigma\lambda_\nu\neq 0,$$

and where

$$(119) \qquad \rho_\nu=e_p(\sigma\lambda_\nu).$$

Let θ be a characteristic root of $L(t,R,X,q)$, that is the reciprocal of a zero or a pole of $L(t,R,X,q)$. If θ does not disappear in the left hand side of (117), then it must be a characteristic root of the right hand side of (117), and equation (108) must be satisfied. Suppose that θ disappears in the left hand side of (117); then θ^{p^μ} must be a characteristic root of $L(t,R,X,q^{p^\mu})$ whence, for every γ such that $\gamma^{p^\mu}=1$, the number $\gamma\theta$ is a characteristic root of $L(t^{p^\mu},R,X,q^{p^\mu})$. If one of the roots $\gamma\theta$ does not disappear in (117), then by our previous argument equation (108) is verified. Suppose that all roots $\gamma\theta$ disappear; we deduce that the function $L(t,R,X,q)^{p^\mu}$ has at least p^μ distinct characteristic roots, in fact the roots $\gamma\theta$; as these roots are all distinct, we deduce that the function $L(t,R,X,q)$ has p^μ distinct characteristic rots, and this is impossible for μ large enough, because $L(t,R,X,q)$ is a rational function, hence with a bounded number of characteristic roots. This proves (108) and Theorem 7.

Remarks. It may be useful to give some comments on Theorem 7.

If X is a complete non-singular curve, then by (i) of Theorem A the function $L(t,R,X)$ is a polynomial; now we have

$$N_{K/Q}L(t,R,X)=Z(t,Y)/Z(t,Y),$$

and it follows immediately that the characteristic roots of $L(t,R,X)$ are also characteristic roots of $Z(t,Y)$. This is in fact the argument used by Weil in his proof of the Riemann hypothesis for Artin L-series of complete non-singular curves. In the general case the function $L(t,R,X)$ is only a rational function, and there may be cancellation of roots in the norm $N_{K/Q}L(t,R,X)$; this is the reason why we had to use the more complicated identity (117). Lastly, equations (109) and (106) suggest that if ω is a characteristic root of $Z(t, Y(\lambda), q^{p^\nu})$ then ω^p is a characteristic root of $Z(t, Y, q^{p^{\nu+1}})$, so that we are led to the conjecture that if θ is a characteristic root of $L(t, R, X)$ then

$$\theta = \gamma\omega, \tag{120}$$

where ω is a characteristic root of $Z(t, Y, q)$ and γ is a p^μ-th root of unity.

We may ask how far the proof of Theorem 7 can be extended to other cases, not necessarily of the Artin-Schreier type. In a paper which is now in course of publication we have generalized our identity (117) to every L-function corresponding to a separable, normal, unramified Galois covering over a finite field, and in particular we can prove that the Corollary to Theorem 7 is true for any *rational* L-function corresponding to a covering of the previous type. As this generalization depends on the general theory of group varieties, and thus is outside the general plan of this paper, we hope to return on this aspect of the theory of Artin L-functions in another occasion.

CAMBRIDGE, ENGLAND.

REFERENCES.

[1] B. Dwork, "On the rationality of the Zeta function of an algebraic variety," *American Journal of Mathematics*, vol. 82 (1960), pp. 631-648.

[2] ———, *On the Zeta function of a hypersurface*, Publications Mathematiques Institut des Hautes Etudes Scientifiques 12 (1962).

[3] ———, "On the Zeta function of a hypersurface II," *Annals of Mathematics*, vol. 80 (1964), pp. 227-239.

[4] H. Hasse, "Theorie der relativ-zyklischen algebraischen Funktionenkörpern, insbesondere bei endlichem Konstantenkörper," *Crelles Journal für Mathematik*, vol. 172 (1934), pp. 37-54.

[5] S. Lang, "Unramified class field theory over function fields in several variables," *Annals of Mathematics*, vol. 64 (1956), pp. 285-325.

[6] ———, *Abelian varieties*, Interscience Tracts no. 7, New York (1959).

[7] J. P. Serre, "Endomorphismes complètement continus des espaces de Banach p-adiques," *Publications Mathématiques Institut des Hautes Etudes Scientifiques*, vol. 12, Paris (1962).

[8] ———, *Groupes algébriques et corps de classes*, Paris, Hermann (1959).

[9] ———, *Corps locaux*, Paris, Hermann (1962).

[10] ———, "Algèbre et géométrie," *Annuaire du Collège de France*, vol. 62 (1962), pp. 47-51.

[11] A. Weil, "On some exponential sums," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 34 (1948), pp. 204-207.

[12] ———, *Sur les courbes algébriques et les variétés qui s'en déduisent*, Paris, Hermann (1948).

THE STRUCTURE OF THE SYMPLECTIC GROUP OVER A VALUATION RING.

By Carl R. Riehm.*

The symplectic group was first investigated by Abel, and since that time questions relating to its simplicity have been dealt with by Jordan [5], Dickson [2] and Dieudonné [3]. With but three exceptional cases it has been found that the only proper normal subgroup of a symplectic group is $\{\pm I\}$—see [1], for example.

More recently Klingenberg [6] effected a significant generalization of these results by extending them to arbitrary commutative local rings (with suitable residue class field). In doing so however, he assumed that the underlying alternating form is unimodular, i.e. has unit discriminant. He was able to prove in this case that every normal subgroup of the symplectic group is a congruence subgroup in the usual sense.

In this paper we shall generalize these latter results by dropping the requirement that the discriminant be a unit, and requiring only that it be non-zero (i.e. that the alternating form be non-degenerate). To do this it is necessary to restrict the ring to be a valuation ring (with residue class field $\neq \boldsymbol{F}_3$ and not of characteristic 2, the same assumptions made by Klingenberg). *In this situation it turns out that there are normal subgroups which are not congruence subgroups.*

In order to regain a complete description of the normal subgroups, we generalize the concept of congruence subgroup by allowing the congruence ideal to vary from entry to entry in the matrix. (A precise definition is given in § 4). We are then able to prove the principal result of this paper: *every congruence subgroup in the new sense is a normal subgroup of the symplectic group, and conversely every normal subgroup is a congruence group in the new sense.* These and other theorems are summarized in the *Structure Theorem for $Sp(L)$* at the end of the paper.

We have avoided completely the use of matrices. Instead we describe congruence groups by their effect on "invariant lattices" (see § 2). In an announcement [7] of the present paper, the reader will find a matricial interpretation of our results.

Received July 27, 1964.
Revised May 28, 1965.

* Research partially supported under National Science Foundation grants GP-1656 and GP-3986.

§ 1. Basic Properties of Lattices.

From now on $\mathfrak{o}$ will be a valuation ring, that is an integral domain which has the property that given any two of its elements, one is a divisor of the other. Easily verified equivalent conditions are (i) the principal ideals of $\mathfrak{o}$ are totally ordered by inclusion, or (ii) all ideals of $\mathfrak{o}$ are totally ordered by inclusion. An important propery of $\mathfrak{o}$ is that if an ideal is generated by a given finite set of elements, then one of these elements already generates the ideal. In particular any finitely generated ideal is principal, and hence also invertible. Denote by F the field of quotients of $\mathfrak{o}$. If $\alpha \in F$, the (fractional) principal ideal generated by α is denoted by $[\alpha]$.

It is easily verified that $\mathfrak{o}$ is a local ring; denote its unique maximal ideal by $\mathfrak{p}$.

Let L be a free module of finite rank over $\mathfrak{o}$. Then $L/\mathfrak{p}L$ can be viewed as a vector space over $\mathfrak{o}/\mathfrak{p}$. We say that a set of vectors of L are linearly independent mod $\mathfrak{p}$ if their images in $L/\mathfrak{p}L$ are linearly independent.

The following properties can be verified directly.

(i) A set of vectors in L can be extended to a basis of L if and only if they are independent mod $\mathfrak{p}$.

(ii) Any direct summand of L is free.

A *lattice* L is a free $\mathfrak{o}$-module of finite rank, endowed with a non-degenerate alternating bilinear form. Thus there is a bilinear mapping of $L \times L$ into $\mathfrak{o}$, denoted by $(X, Y) \to XY$, such that $X^2 = 0$ for all X in L (whence $XY = -YX$ for all X and Y), and such that $XL = 0$ only for $X = 0$. The rank of L is denoted by $\dim L$. Although it will not be mentioned explicitly very often, it is convenient to think of L as being imbedded in the vector space $FL = F \otimes_{\mathfrak{o}} L$ over F. The alternating form has a canonical extension to FL.

The *symplectic group* $Sp(L)$ is the group of those automorphisms σ of L such that $(\sigma X)(\sigma Y) = XY$ for all X and Y in L. We may suppose that $Sp(L) \subseteq Sp(FL)$.

Notice that if non-empty subsets U and V of a lattice L span finitely generated modules, then the ideal generated by UV is finitely generated, hence principal; more precisely the ideal generated by UV is of the form $[XY]$ for suitable $X \in U$ and $Y \in V$. In particular $LL = [XY]$ for suitable X and Y in L; for example X and Y can be chosen as vectors in a given basis for L.

A *canonical basis* for a lattice L is a basis $\{E_1, \cdots, E_{2n}\}$ with the following properties:

(i) *if* $i < j$, *then* $E_iE_j = 0$ *unless* $i = 2k - 1$, $j = 2k$ *for some* k, $1 \leqq k \leqq n$.

(ii) *if* $\alpha_i = E_{2i-1}E_{2i}$ *for* $i = 1, \cdots, n$, *then*

$$\alpha_1 \mid \alpha_2 \mid \cdots \mid \alpha_n.$$

(Here $\mid$ stands for "divides"). Note that non-degeneracy implies that α_i is non-zero for all i.

Theorem 1.1. *A lattice L has a canonical basis.*

Proof. Let $D_1, \cdots, D_m$ be a basis for L. Renumbering if necessary we may suppose that $[D_1D_2] = LL$. We may assume that $m > 2$. For $j = 3, \cdots, m$, define α_j and β_j in $\mathfrak{o}$ by

$$D_1D_j = \beta_j(D_1D_2), \qquad D_jD_2 = \alpha_j(D_1D_2),$$

and put $D_j' = D_j - \alpha_jD_1 - \beta_jD_2$. Let M (resp. N) be the module spanned by $\{D_1, D_2\}$ (resp. $\{D_3', \cdots, D_m'\}$). Then it is easy to check that $L = M \perp N$ (orthogonal direct sum). By induction on the dimension, N has a canonical basis $\{E_3, \cdots, E_{2n}\}$. Let $E_1 = D_1$ and $E_2 = D_2$. Since $[E_1E_2] = LL \supseteq NN$, $\{E_1, \cdots, E_{2n}\}$ is a canonical basis for L.

A vector A in a lattice L is called *maximal* (*in* L) if $A = \alpha B$ with B in L implies α is a unit. This is equivalent to saying that if A is expressed in terms of a basis of L, at least one of the coefficients is a unit, or alternatively that the image of A in $L/\mathfrak{p}L$ is non-zero. It is easy to see that any vector of L is a multiple of some maximal vector of L.

We call a non-zero lattice L *modular* if $AL = LL$ for each maximal vector A in L.

If $\mathbf{X}$ is a subset of a lattice L, then $\mathbf{X}^0$ is the submodule of L consisting of those vectors A orthogonal to $\mathbf{X}$: $A\mathbf{X} = 0$. A sublattice M of L is said to *split* L if $L = M \perp M^0$.

Lemma 1.2. *Let M be a modular sublattice of the lattice L with $ML = MM$. Then M splits L.*

Proof. Let A be an maximal vector of M, and suppose that $A = \alpha B$ with B in L and α in $\mathfrak{o}$. We can find A' in M such that $[AA'] = MM$. Then $MM = \alpha[BA'] \subseteq \alpha ML$ and it follows that α is a unit since ML is invertible. Thus any maximal vector of M is maximal in L.

Let $\{E_1, \cdots, E_{2n}\}$ be a canonical basis of M; then these vectors are independent mod $\mathfrak{p}$ (as vectors of L). For if $\sum \alpha_jE_j$ is in $\mathfrak{p}L$, it is not maximal in L, hence not maximal in M, hence in $\mathfrak{p}M$. Thus we can extend this canonical basis of M to a basis $\{E_1, \cdots, E_{2m}\}$ of L.

Now define α_{ij} for $1 \leqq i \leqq 2n$ and $2n+1 \leqq j \leqq 2m$ as follows: $E_iE_j = \alpha_{i+1,j}E_iE_{i+1}$ if i is odd, $E_jE_i = \alpha_{i-1,j}E_{i-1}E_i$ if i is even. From $[E_{2k-1}E_{2k}] = MM = ML$ for $k = 1, \cdots, n$ it follows that each α_{ij} is in $\mathfrak{o}$. Let N be the submodule spanned by $E_j - \sum_{i=1}^{2n} \alpha_{ij}E_i$ for $j = 2n+1, \cdots, 2m$. Clearly $L = M \oplus N$ and $MN = 0$ whence $N = M^0$ and $L = M \perp M^0$ as required.

Now if $\{E_1, \cdots, E_{2n}\}$ is a canonical basis for a lattice L, each sublattice $\mathfrak{o}E_{2i-1} + \mathfrak{o}E_{2i}$ is modular. If $K \perp M$ is a sublattice of L with K and M modular and $KK = MM$, then $K \perp M$ is also modular. Thus we see that L has at least one splitting

$$L = L_1 \perp \cdots \perp L_t \tag{1.1}$$

where each L_i is modular, and if $L_iL_i = \mathfrak{s}_i$ for $i = 1, \cdots, t$, then each $\mathfrak{s}_i$ is principal and $\mathfrak{s}_1 \supset \mathfrak{s}_2 \supset \cdots \supset \mathfrak{s}_t$ (proper containment). Such a splitting is called a *canonical splitting.*

Let $\mathfrak{a}$ be a non-zero ideal and M a submodule of the lattice L. We define

$$M^{\mathfrak{a}} = \{A \in M : AM \subseteq \mathfrak{a}\}.$$

Clearly $M^{\mathfrak{a}}$ is a submodule of M, and $M^{\mathfrak{a}}M^{\mathfrak{a}} \subseteq \mathfrak{a}$. If $M = J \perp K$ then $M^{\mathfrak{a}} = J^{\mathfrak{a}} \perp K^{\mathfrak{a}}$. If $MM \subseteq \mathfrak{a}$, $M^{\mathfrak{a}} = M$. If M is modular with $\mathfrak{a} \subseteq \mathfrak{b} = MM$, then it is easy to show that $M^{\mathfrak{a}} = \mathfrak{a}\mathfrak{b}^{-1}M$.

It follows that if (1.1) is a canonical splitting for L, and $\mathfrak{s}_i = L_iL_i$ as before, then

$$L^{\mathfrak{a}} = \mathop{\perp}_{1 \leqq i \leqq t} ((\mathfrak{a}\mathfrak{s}_i^{-1}) \cap \mathfrak{o})L_i. \tag{1.2}$$

In particular $L^{\mathfrak{a}}$ is a lattice (i. e. is free) if and only if either $\mathfrak{a}$ is a principal ideal or $LL \subseteq \mathfrak{a}$. Also

$$L^{\mathfrak{a}}L^{\mathfrak{a}} = \mathfrak{a} \Longleftrightarrow \mathfrak{a} = \mathfrak{s}_i \text{ for some } i. \tag{1.3}$$

The lattices $L^{\mathfrak{s}_1}, \cdots, L^{\mathfrak{s}_t}$ are of particular importance and satisfy

$$\begin{gathered} L^{\mathfrak{s}_1} \supset L^{\mathfrak{s}_2} \supset \cdots \supset L^{\mathfrak{s}_t} \\ \mathfrak{s}_1^{-1}L^{\mathfrak{s}_1} \subset \mathfrak{s}_2^{-1}L^{\mathfrak{s}_2} \subset \cdots \subset \mathfrak{s}_t^{-1}L^{\mathfrak{s}_t} \end{gathered} \tag{1.4}$$

where $\subset$ denotes strict inclusion.

THEOREM 1.3. *The quantities* t, $\mathfrak{s}_1, \cdots, \mathfrak{s}_t$ *and* $\dim L_1, \cdots, \dim L_t$ *derived from a canonical splitting* (1.1) *of* L *are invariants of* L*—that is,* *they are the same no matter which canonical splitting of L is used to calculate them.*

Proof. The ideals $\mathfrak{s}_1, \cdots, \mathfrak{s}_t$ are characterized as those ideals $\mathfrak{a}$ satisfying (1.3) and hence they and t are invariant. On the other hand using (1.2), the reader can check that $L^{\mathfrak{s}_i}/(L^{\mathfrak{s}_i})^{\mathfrak{p}\mathfrak{s}_i} \cong L_i/\mathfrak{p}L_i$, whence

$$\dim L_i = \dim_{\mathfrak{o}/\mathfrak{p}}(L^{\mathfrak{s}_i}/(L^{\mathfrak{s}_i})^{\mathfrak{p}\mathfrak{s}_i}).$$

This finishes the proof.

COROLLARY 1.4. *Let $\{E_1, \cdots, E_{2n}\}$ be a canonical basis for the lattice L, and put $\alpha_i = E_{2i-1}E_{2i}$ for $i = 1, \cdots, n$. Then*

$$[\alpha_1] \supseteq [\alpha_2] \supseteq \cdots \supseteq [\alpha_n] \tag{1.5}$$

and these ideals are uniquely determined by L, i.e. any other canonical basis for L gives rise to the same ideals.

Proof. That (1.5) holds is a consequence of the definition of canonical basis. Now each ideal $[\alpha_j]$ is one of $\mathfrak{s}_1, \mathfrak{s}_2, \cdots, \mathfrak{s}_t$ and the number of times an $\mathfrak{s}_i$ appears in (1.5) is equal to $\frac{1}{2}\dim L_i$, as can be seen by constructing the canonical splitting from the given canonical basis in the obvious manner. Then the uniqueness of (1.5) follows from Theorem 1.3.

Throughout the rest of the paper we shall be working with a lattice L. Whenever the symbols t, $\mathfrak{s}_i$, $\dim L_i$ appear, they will be the invariants of L given in the foregoing theorem unless stated otherwise explicitly.

§ 2. Invariant Modules.

A subset $\mathbf{X}$ of FL is called *bounded* if there is a non-zero element ξ of $\mathfrak{o}$ such that $\xi\mathbf{X} \subseteq L$. An $\mathfrak{o}$-submodule $\mathfrak{L}$ of FL is called *invariant* (with respect to $Sp(L)$) if it is bounded and if $\sigma\mathfrak{L} \subseteq \mathfrak{L}$ for all σ in $Sp(L)$. We shall determine all such invariant modules when $\mathfrak{o}$ is an arbitrary valuation ring.

A *pure* vector A of L is one which is in some canonical basis for L; A is pure *of type* i if $AL = \mathfrak{s}_i$.

LEMMA 2.1. *If an invariant module contains a pure vector of type i, then it contains all of $L^{\mathfrak{s}_i}$.*

Proof. Let A be a pure vector of type i contained in the invariant module $\mathfrak{L}$. Choose a canonical basis for L containing A and let B be the vector in this basis with $AB \neq 0$. Then $L^{\mathfrak{s}_i} = (\mathfrak{o}A + \mathfrak{o}B) \perp M$ by Lemma 1.2. Choose a canonical basis for M and let C be any vector in this basis or let $C = B$. Then $L = (\mathfrak{o}(A + C) + \mathfrak{o}B) \perp K$ by Lemma 1.2. Choose a canonical basis for K and adjoin $A + C$ and B to it; this yields a canonical basis for L. By the invariant properties of canonical bases (Corollary 1.4) there is a map σ in $Sp(L)$ such that $\sigma A = A + C$. Thus $C = (A + C) - A \in \mathfrak{L}$, and so $L^{\mathfrak{s}_i} \subseteq \mathfrak{L}$ as required.

THEOREM 2.2. *If* $\mathfrak{L}$ *is an invariant module with respect to* $Sp(L)$, *there are ideals* $\mathfrak{g}_1, \cdots, \mathfrak{g}_t$ *such that*

$$\mathfrak{L} = \mathfrak{g}_1 L^{\mathfrak{s}_1} + \mathfrak{g}_2 L^{\mathfrak{s}_2} + \cdots + \mathfrak{g}_t L^{\mathfrak{s}_t}, \tag{2.1}$$

$$\mathfrak{g}_1 \subseteq \mathfrak{g}_2 \subseteq \cdots \subseteq \mathfrak{g}_t, \; \mathfrak{s}_1\mathfrak{g}_1 \supseteq \mathfrak{s}_2\mathfrak{g}_2 \supseteq \cdots \supseteq \mathfrak{s}_t\mathfrak{g}_t. \tag{2.2}$$

These ideals are uniquely determined and are in fact given by

$$\mathfrak{g}_j = \mathfrak{s}_j^{-1}(\mathfrak{L}L^{\mathfrak{s}_j}). \tag{2.3}$$

If $L = L_1 \perp \cdots \perp L_t$ *is an arbitrary canonical splitting, then*

$$\mathfrak{L} = (\mathfrak{g}_1 L_1) \perp (\mathfrak{g}_2 L_2) \perp \cdots \perp (\mathfrak{g}_t L_t). \tag{2.4}$$

Remarks. It is easy to check that $L^{\mathfrak{a}}$ is invariant for any non-zero ideal $\mathfrak{a}$, and hence if $\mathfrak{g}_1, \cdots, \mathfrak{g}_t$ are any ideals, the module defined by (2.1) is invariant. Moreover if these ideals also satisfy (2.2), it follows from the theorem that the module defined by (2.4) is invariant.

Since the ideals $\mathfrak{g}_1, \cdots, \mathfrak{g}_t$ characterize $\mathfrak{L}$ completely, we shall often write $\mathfrak{L} = (\mathfrak{g}_1, \cdots, \mathfrak{g}_t)$. Notice that it follows from the theorem that a non-zero invariant module contains a non-zero vector from every line of FL.

Proof of Theorem 2.2. Choose a canonical splitting (1.1) and fix it for the rest of the proof. Let $\mathscr{L}$ be the set of all modules in FL of the form $\sum_{j=1}^{t} \mathfrak{a}_j L_j$ where $\mathfrak{a}_1, \cdots, \mathfrak{a}_t$ are ideals of $\mathfrak{o}$ (possibly fractional). Since $\mathfrak{L}$ is bounded some member of $\mathscr{L}$ contains $\mathfrak{L}$. Since $\mathscr{L}$ is clearly closed under intersections, the intersection of all members of $\mathscr{L}$ which contain $\mathfrak{L}$ is of the form $\sum \mathfrak{g}_j L_j$. We shall show that (2.1) and (2.4) hold for this choice of the ideals $\mathfrak{g}_1, \cdots, \mathfrak{g}_t$.

Trivially $\mathfrak{L} \subseteq \sum \mathfrak{g}_j L_j$. Suppose $\mathfrak{g}_i \neq 0$ for some i, and select any non-zero element γ from $\mathfrak{g}_i$. Then there exists a vector A in $\mathfrak{L}$ with the following property: if we write $A = \sum \alpha_j E_j$ with E_j maximal in L_j for each j, then $\gamma \in [\alpha_i]$. For otherwise we would have

$$\mathfrak{L} \subseteq \gamma\mathfrak{p}L_i + \sum_{j \neq i} \mathfrak{g}_j L_j \subset \sum \mathfrak{g}_j L_j$$

which is a contradiction. Now let B be a vector of L_i such that E_i and B are independent mod $\mathfrak{p}$; then we can find $\sigma \in Sp(L)$ such that $\sigma \mid L_i^0$ is the identity and $\sigma E_i = B$. The vector $C = E_i - B$ is a pure vector of type i, and $\alpha_i C = A - \sigma A \in \mathfrak{L}$. Thus $C \in \alpha_i^{-1}\mathfrak{L}$ and so $\gamma L^{\mathfrak{s}_i} \subseteq \alpha_i L^{\mathfrak{s}_i} \subseteq \mathfrak{L}$ by Lemma 2.1. Hence $\mathfrak{g}_i L_i \subseteq \mathfrak{g}_i L^{\mathfrak{s}_i} \subseteq \mathfrak{L}$, and so $\mathfrak{L} = \sum \mathfrak{g}_j L_j = \sum \mathfrak{g}_j L^{\mathfrak{s}_j}$ as stated.

Since $\mathfrak{g}_i L^{\mathfrak{s}_i} \subseteq \mathfrak{L}$ for each i, we have by (1.2)

$$\sum_j ((\mathfrak{s}_i \mathfrak{s}_j^{-1}) \cap \mathfrak{o}) \mathfrak{g}_i L_j \subseteq \sum_j \mathfrak{g}_j L_j$$

from which the relations (2.2) follow. The formula (2.3) can be proved easily by direct calculation of $\mathfrak{L} L^{\mathfrak{s}_j}$ using (2.4), (1.2) and (2.2).

If $\mathbf{X}$ is a bounded subset of FL, we denote by $\mathfrak{I}(\mathbf{X})$ the smallest invariant module containing $\mathbf{X}$—its existence is easy to prove. It is the intersection of all invariant modules containing $\mathbf{X}$ and is also given as the set of all finite sums

$$\mathfrak{I}(\mathbf{X}) = \{\sum \alpha_k \sigma_k X_k \colon \alpha_k \in \mathfrak{o}, \sigma_k \in Sp(L), X_k \in \mathbf{X}\} \tag{2.5}$$

We note that $\mathfrak{I}(\mathfrak{I}(\mathbf{X})) = \mathfrak{I}(\mathbf{X})$.

LEMMA 2.3. *Let* $\mathbf{X}$ *be a bounded subset of* FL *and let* $\mathfrak{g}_j = \mathfrak{s}_j^{-1}(\mathbf{X} L^{\mathfrak{s}_j})$ *for* $j = 1, \cdots, t$. *Then* $\mathfrak{I}(\mathbf{X}) = (\mathfrak{g}_1, \cdots, \mathfrak{g}_t)$.

Proof. The conditions (2.2) follow easily from (1.4) and so $\mathfrak{L} = (\mathfrak{g}_1, \cdots, \mathfrak{g}_t)$ is invariant. Now let $\mathfrak{I}(\mathbf{X}) = (\mathfrak{g}_1', \cdots, \mathfrak{g}_t')$. From (2.3), $\mathfrak{g}_j \subseteq \mathfrak{g}_j'$, whence $\mathfrak{L} \subseteq \mathfrak{I}(\mathbf{X})$. Conversely let $X \in \mathbf{X}$ and suppose that $X = \sum_{j=1}^{t} \xi_j E_j$ where E_j is maximal in L_j (and (1.1) is a canonical splitting). Choose E_j' in L_j so that $[E_j E_j'] = \mathfrak{s}_j$. Then $\mathfrak{s}_j \mathfrak{g}_j \supseteq [X E_j'] = \xi_j \mathfrak{s}_j$ whence $\xi_j \in \mathfrak{g}_j$. Thus $\mathbf{X} \subseteq \mathfrak{L}$ whence $\mathfrak{I}(\mathbf{X}) \subseteq \mathfrak{L}$ as required.

§ 3. Orders.

We shall define the order $o(G)$ for any non-empty subset G of $Sp(L)$, in particular for an element or subgroup. From now on we must assume that $\operatorname{char} \mathfrak{o}/\mathfrak{p} \neq 2$, i. e. 2 is a unit of $\mathfrak{o}$.

The mappings $X \to X$ and $X \to -X$ of L are both in $Sp(L)$ and will be denoted by 1_L and -1_L respectively.

Let $\sigma \in Sp(L)$; we define

$$o_i(\sigma) = \bigcap_{\delta = \pm 1} \mathfrak{I}((\sigma - \delta_L) L^{\mathfrak{s}_i}), \qquad i = 1, \cdots, t.$$

Clearly $o_i(\sigma)$ is an invariant module and

$$(3.1) \qquad o_i(\sigma) \subseteq L^{\mathfrak{s}_i}.$$

Now if $X \in L^{\mathfrak{s}_i}$, then $X = \frac{1}{2}((\sigma + 1_L)X - (\sigma - 1_L)X)$ whence

$$(3.2) \qquad \sum_{\delta=\pm 1} \mathfrak{I}((\sigma - \delta_L)L^{\mathfrak{s}_i}) = L^{\mathfrak{s}_i}, \qquad i = 1, \cdots, t.$$

Thus one of the two summands of (3.2) contains a pure vector of type i, whence one of them is $L^{\mathfrak{s}_i}$ by Lemma 2.1 (and so one of them always contains the other); for $1 \leqq i \leqq t$ we have

$$(3.3) \qquad o_i(\sigma) = \min_{\delta=\pm 1} \{\mathfrak{I}((\sigma - \delta_L)L^{\mathfrak{s}_i})\}, L^{\mathfrak{s}_i} = \max_{\delta=\pm 1} \{\mathfrak{I}((\sigma - \delta_L)L^{\mathfrak{s}_i})\}.$$

It is clear that $o_i(\sigma) = 0$ if and only if $\sigma = \pm 1_L$. And if $\sigma \neq \pm 1_L$, then $o_i(\sigma)$ spans FL.

We now define the *order* $o(\sigma)$ to be the ordered t-tuplet $\{o_1(\sigma), \cdots, o_t(\sigma)\}$ of invariant modules. We often write $o(\sigma) = \{o_i(\sigma)\}$. When convenient we also write $o(\sigma) = \{o_{ij}(\sigma)\}$ where the ideals $o_{ij}(\sigma)$ for $i, j = 1, \cdots, t$ are defined by $o_i(\sigma) = (o_{i1}(\sigma), \cdots, o_{it}(\sigma))$. One can show, using Lemma 2.3 and (3.3), that

$$(3.4) \qquad \mathfrak{s}_j o_{ij}(\sigma) = \min_{\delta=\pm 1} \{((\sigma - \delta_L)L^{\mathfrak{s}_i})L^{\mathfrak{s}_j}\}.$$

Thus $o_{ij}(\sigma)$ can be defined directly, i.e. without reference to § 2; however that section is necessary to develop the properties of $o(\sigma)$. The formula (3.4) also shows that $o_{ij}(\sigma)$ is principal, since $(\sigma - \delta_L)L^{\mathfrak{s}_i}$ and $L^{\mathfrak{s}_j}$ are finitely generated modules. Thus $o_i(\sigma)$ is actually a lattice (see, for example, (2.4)).

Let us temporarily put $o_i(\sigma) = \mathfrak{L}_i$. Then we have by (3.1) and (1.4)

$$(3.5) \qquad \mathfrak{L}_i \subseteq L^{\mathfrak{s}_i} \qquad i = 1, \cdots, t.$$

$$(3.6) \qquad \mathfrak{L}_1 \supseteq \mathfrak{L}_2 \supseteq \cdots \supseteq \mathfrak{L}_t$$

$$(3.7) \qquad \mathfrak{s}_1^{-1}\mathfrak{L}_1 \subseteq \mathfrak{s}_2^{-1}\mathfrak{L}_2 \subseteq \cdots \subseteq \mathfrak{s}_t^{-1}\mathfrak{L}_t.$$

If we put $o_{ij}(\sigma) = \mathfrak{g}_{ij}$, then these conditions can be restated as

$$(3.8) \qquad \mathfrak{g}_{ij} \subseteq (\mathfrak{s}_i \mathfrak{s}_j^{-1}) \cap \mathfrak{o} \qquad i, j = 1, \cdots, t$$

$$(3.9) \qquad \mathfrak{g}_{1j} \supseteq \mathfrak{g}_{2j} \supseteq \cdots \supseteq \mathfrak{g}_{tj}, \mathfrak{g}_{i1} \subseteq \mathfrak{g}_{i2} \subseteq \cdots \subseteq \mathfrak{g}_{it}$$

$$(3.10) \qquad \mathfrak{s}_1^{-1}\mathfrak{g}_{1j} \subseteq \mathfrak{s}_2^{-1}\mathfrak{g}_{2j} \subseteq \cdots \subseteq \mathfrak{s}_t^{-1}\mathfrak{g}_{tj}, \mathfrak{s}_1\mathfrak{g}_{i1} \supseteq \mathfrak{s}_2\mathfrak{g}_{i2} \supseteq \cdots \supseteq \mathfrak{s}_t\mathfrak{g}_{it}$$

where the second members of (3.9) and (3.10) derive from (2.2).

LEMMA 3.1. *Suppose that* $\sigma \in Sp(L)$ *and that* $o_i(\sigma) = \mathfrak{I}((\sigma - 1_L)L^{\mathfrak{s}_i})$ *and* $o_j(\sigma) = \mathfrak{I}((\sigma + 1_L)L^{\mathfrak{s}_j})$. *Then*

$$o_{ij}(\sigma) = (\mathfrak{s}_i\mathfrak{s}_j^{-1}) \cap \mathfrak{o} \text{ and } o_{ji}(\sigma) = (\mathfrak{s}_j\mathfrak{s}_i^{-1}) \cap \mathfrak{o}.$$

Proof. We shall suppose that $i \leqq j$. Then $L^{\mathfrak{s}_j} = \mathfrak{I}((\sigma - 1_L)L^{\mathfrak{s}_j}) \subseteq o_i(\sigma)$ whence $o_{ij}(\sigma) = \mathfrak{o}$ as stated. In a similar manner $L^{\mathfrak{s}_i} \subseteq \mathfrak{s}_j^{-1}\mathfrak{s}_i o_j(\sigma)$ whence whence $o_{ji}(\sigma) = \mathfrak{s}_j\mathfrak{s}_i^{-1}$ as required.

Under the conditions of the lemma we have, with $\mathfrak{g}_{ij} = o_{ij}(\sigma)$,

$$\mathfrak{s}_i\mathfrak{g}_{ji} = \mathfrak{s}_j\mathfrak{g}_{ij}. \tag{3.11}$$

On the other hand if $\delta = \pm 1$, then $(\sigma X - \delta X)Y = \delta(\sigma X - \delta Y)\sigma X$ whence $((\sigma - \delta_L)L^{\mathfrak{s}_j})L^{\mathfrak{s}_i} = ((\sigma - \delta_L)L^{\mathfrak{s}_i})L^{\mathfrak{s}_j}$ whence (3.11) holds in all cases by (3.4).

Definition. A set of ideals $\{\mathfrak{g}_{ij}\}$, $(i, j = 1, \cdots, t)$, *satisfying* (3.8), (3.9), (3.10) *and* (3.11) *is called a tableau.*

By what we have proved above, $o(\sigma) = \{o_{ij}(\sigma)\}$ is a tableau for $\sigma \in Sp(L)$.

If $\{\mathfrak{g}_{ij}\}$ is a tableau and we put $\mathfrak{L}_i = (\mathfrak{g}_{i1}, \cdots, \mathfrak{g}_{it})$ for $i = 1, \cdots, t$, then it is easy to see that $\mathfrak{L}_1, \cdots, \mathfrak{L}_t$ are invariant modules satisfying (3.5), (3.6) and (3.7).

We order the set of all tableaus by defining $\{\mathfrak{g}_{ij}\} \subseteq \{\mathfrak{h}_{ij}\}$ if $\mathfrak{g}_{ij} \subseteq \mathfrak{h}_{ij}$ for and i and j. With this partial ordering the set of all tableaus has a largest member and a smallest. The smallest is the *zero tableau* given by $\mathfrak{g}_{ij} = 0$ for all i and j, while the largest is the *maximum tableau* given by $\mathfrak{g}_{ij} = (\mathfrak{s}_i\mathfrak{s}_j^{-1}) \cap \mathfrak{o}$ for all i and j (cf. (3.8)); it is the unique tableau for which $\mathfrak{g}_{ii} = \mathfrak{o}$ for all i.

Definition. Let G *be any non-empty subset of* $Sp(L)$. *Define* $o_{ij}(G) = \sum_{\sigma \in G} o_{ij}(\sigma)$ *and* $o_i(G) = (o_{i1}(G), \cdots, o_{it}(G))$ *for* $i, j = 1, \cdots, t$. *We let* $o(G) = \{o_{ij}(G)\}$ *be the order of* G.

Clearly $\{o_{ij}(G)\}$ is a tableau.

LEMMA 3.2. *Let* $\sigma, \rho \in Sp(L)$. *If* $\delta, \epsilon = \pm 1$, *then*

$$\mathfrak{I}((\sigma\rho - (\delta\epsilon)_L)L^{\mathfrak{s}_i}) \subseteq \mathfrak{I}((\sigma - \delta_L)L^{\mathfrak{s}_i}) + \mathfrak{I}((\rho - \epsilon_L)L^{\mathfrak{s}_i})$$

and $o_i(\sigma\rho) \subseteq o_i(\sigma) + o_i(\rho)$ *for* $i = 1, \cdots, t$.

Proof. The last part follows from the first and the definitions, while the first follows from the identity $\sigma\rho - (\delta\epsilon)_L = (\sigma - \delta_L)\rho + \delta_L(\rho - \epsilon_L)$.

§ 4. Congruence Subgroups of $Sp(L)$.

Let $\{\mathfrak{g}_{ij}\}$ be a tableau. The set of all σ in $Sp(L)$ with $o(\sigma) \subseteq \{\mathfrak{g}_{ij}\}$ is called the *general congruence subgroup of* $Sp(L)$ *corresponding to* $\{\mathfrak{g}_{ij}\}$. It is denoted by $GSp(L;\{\mathfrak{g}_{ij}\})$, or simply GSp when the lattice and tableau are clearly determined by the context.

Let $\mathfrak{L}_i = (\mathfrak{g}_{i1}, \cdots, \mathfrak{g}_{it})$ for $i = 1, \cdots, t$ (for notation see the remark preceding the proof of Theorem 2.2). Then GSp is also characterized as the set of σ such that $o_i(\sigma) \subseteq \mathfrak{L}_i$ for all i, whence GSp is closed under multiplication by Lemma 3.2. Moreover if $\delta = \pm 1$, then $(\sigma^{-1} - \delta_L)L^{\mathfrak{s}_i} = (\sigma - \delta_L)L^{\mathfrak{s}_i}$ since $L^{\mathfrak{s}_i} = \sigma L^{\mathfrak{s}_i}$, whence $o(\sigma^{-1}) = o(\sigma)$, whence GSp is in fact a subgroup of $Sp(L)$. Also we have

$$\rho\tau\rho^{-1} - \delta_L = \rho(\sigma - \delta_L)\rho^{-1}; \ \rho, \sigma \in Sp(L), \delta = \pm 1, \tag{4.1}$$

and it follows that GSp is a normal subgroup of $Sp(L)$.

The *special congruence subgroup of* $Sp(L)$ *corresponding to* $\{\mathfrak{g}_{ij}\}$ is defined as the set of all $\sigma \in Sp(L)$ such that $(\sigma - 1_L)L^{\mathfrak{s}_i} \subseteq \mathfrak{L}_i$ for $i = 1, \cdots, t$—equivalently, such that $\mathfrak{l}((\sigma - 1_L)L^{\mathfrak{s}_i}) \subseteq \mathfrak{L}_i$ for all i, or such that $((\sigma - 1_L)L^{\mathfrak{s}_i})L^{\mathfrak{s}_j} \subseteq \mathfrak{s}_j\mathfrak{g}_{ij}$ for $i, j = 1, \cdots, t$. It is denoted by $SSp(L;\{\mathfrak{g}_{ij}\})$, or simply by SSp when the context permits. Clearly $SSp(L;\{\mathfrak{g}_{ij}\}) \subseteq GSp(L;\{\mathfrak{g}_{ij}\})$. As above one shows that SSp is a normal subgroup of $Sp(L)$, and hence *a fortiori* a normal subgroup of GSp.

If $t = 1$, our notions of order and of the special and general congruence subgroups coincide with those of Klingenberg in [6].

Notice that $GSp = SSp = Sp(L)$ if $\{\mathfrak{g}_{ij}\}$ is the maximum tableau, while $SSp = \{1_L\}$ and $GSp = \{\pm 1_L\}$ if $\{\mathfrak{g}_{ij}\}$ is the zero tableau.

In the next theorem Z_2 stands for the group of two elements. If $L = M \perp N$, we shall also employ the common notation $\sigma \perp \rho$ for that mapping of $Sp(L)$ whose restriction to M (resp. N) is $\sigma \in Sp(M)$ (resp. $\rho \in Sp(N)$).

THEOREM 4.1. *Let* $\{\mathfrak{g}_{ij}\}$ *be a tableau. The factor group*

$$GSp(L;\{\mathfrak{g}_{ij}\})/SSp(L;\{\mathfrak{g}_{ij}\})$$

is isomorphic to the direct sum of s *copies of* Z_2, *where* $s \leqq t$. *The integer* s *is* 0 *if* $\{\mathfrak{g}_{ij}\}$ *is the maximum tableau, while otherwise* $s - 1$ *is the number of integers* k, $1 \leqq k \leqq t - 1$, *such that* (i) $\mathfrak{g}_{kk} \subset \mathfrak{o} = \mathfrak{g}_{k,k+1}$, (ii) *there exists* $\kappa > k$ *such that* $\mathfrak{g}_{\kappa\kappa} \subset \mathfrak{o}$. *Finally every subgroup of* GSp *which contains* SSp *is a normal subgroup of* $Sp(L)$.

Proof. We may assume that $\{\mathfrak{g}_{ij}\}$ is not the maximum tableau. Then the set I of integers i such that $\mathfrak{g}_{ii} \subset \mathfrak{o}$ is non-empty. Thus

$$\mathfrak{L}_i = (\mathfrak{g}_{i1}, \cdots, \mathfrak{g}_{it}) \subset L^{\mathfrak{s}_i}$$

for all i in I and it follows from the definition and properties of order that if $\sigma \in GSp$, then $(\sigma - \delta_L)L^{\mathfrak{s}_i} \subseteq \mathfrak{L}_i$ for one and only one choice of $\delta = \pm 1$; we define $\Delta_i(\sigma)$ to be this value of δ. We define $\Delta\colon GSp \to (Z_2)^I$ by $\Delta(\sigma) = (\Delta_i(\sigma))_{i \in I}$. By Lemma 3.2, Δ_i is a homomorphism for each $i \in I$, and therefore so is Δ. It follows immediately from the definition of SSp that the kernel of Δ is exactly SSp. Thus we need only find the image to prove the theorem.

Suppose that $\Delta_i(\sigma) = -\Delta_j(\sigma)$ with $i < j$. Let $k \in I$ be the largest integer $< j$ such that $\Delta_i(\sigma) = \Delta_k(\sigma)$. If $k+1 \notin I$, we have $\mathfrak{o} = \mathfrak{g}_{k+1,k+1} \subseteq \mathfrak{g}_{k,k+1}$. If $k+1 \in I$, then $\Delta_{k+1}(\sigma) = -\Delta_k(\sigma)$ whence

$$o_{k,k+1}(\sigma) = \mathfrak{o} \subseteq \mathfrak{g}_{k,k+1}$$

by Lemma 3.1. Thus in either case there is an integer k, $i \leqq k < j$, such that (i) and (ii) hold.

Let K stand for the set of integers consisting of the largest integer in I, and those k satisfying (i) and (ii). The foregoing considerations show that K determines a partition of I into s subsets (each subset has a member of K as its largest element) with the property that if i and j are in the same subset, then $\Delta_i(\sigma) = \Delta_j(\sigma)$ for each σ in GSp. Thus to show that GSp/SSp has the form stated, it suffices to find, for each k in K, a mapping ϵ_k in GSp such that $\Delta_i(\epsilon_k)$ is -1 if $i \leqq k$ and is 1 if $i > k$.

Take a canonical splitting (1.1) and define for each $k \in K$

$$L(k) = L_1 \perp \cdots \perp L_k, \tag{4.2}$$

and $\epsilon_k = -1_{L(k)} \perp 1_{L(k)^0}$. Trivially $\epsilon_k \in Sp(L)$. If $k < i \leqq t$, we have $(\epsilon_k - 1_L)L^{\mathfrak{s}_i} = L(k)^{\mathfrak{s}_i} \subseteq \mathfrak{s}_i\mathfrak{s}_k^{-1}L^{\mathfrak{s}_k}$. But $\mathfrak{g}_{k+1,k} = \mathfrak{s}_{k+1}\mathfrak{s}_k^{-1}$ by (3.11) and condition (i), whence $\mathfrak{s}_i\mathfrak{s}_k^{-1}L^{\mathfrak{s}_k} \subseteq \mathfrak{s}_i\mathfrak{s}_{k+1}^{-1}\mathfrak{L}_{k+1} \subseteq \mathfrak{L}_i$ by (3.7). Similarly $(\epsilon_k + 1_L)L^{\mathfrak{s}_i} \subseteq \mathfrak{L}_i$ if $i \leqq k$. It follows that ϵ_k has the required properties.

Now let G be a group between SSp and GSp. If $\rho \in Sp(L)$ and $\sigma \in G$, it follows from (4.1) that $\Delta_i(\rho\sigma\rho^{-1}) = \Delta_i(\sigma)$ for all i in I, whence $\Delta(\rho G\rho^{-1}) = \Delta(G)$. Thus $\rho G\rho^{-1} = G$ for all ρ in $Sp(L)$ by the 1-1 correspondence between the subgroup of GSp containing SSp, and the subgroups of $\Delta(GSp)$.

We have in fact virtually proved the following stronger form of Theorem 4.1.

THEOREM 4.2. *Let* $\{\mathfrak{g}_{ij}\}$ *be a tableau which is not maximal and let* K *consists of the largest integer* i *with* $\mathfrak{g}_{ii} \neq \mathfrak{o}$ *and those integers* k *satisfying* (i) *and* (ii) *of Theorem* 4.1. *Fix a canonical splitting* (1.1) *and define* $L(k)$ *for* $k \in K$ *by* (4.2). *Let* E *be the subgroup of* $Sp(L)$ *generated by the* s *mappings*

$$\epsilon_k = -1_{L(k)} \perp 1_{L(k)^\circ}, \quad k \in K.$$

Then GSp *is the semi-direct product of* SSp *by* E. *And any normal subgroup of* GSp *which contains* SSp *is normal in* $Sp(L)$, *and is the semi-direct product of* SSp *by a subgroup of* E.

Proof. We have shown above that $\mathrm{E} \subset GSp$ and that $\Delta(\mathrm{E}) = \Delta(GSp)$. Thus $\Delta \mid \mathrm{E}$ is an isomorphism, and the theorem follows by Theorem 6.5.3 in [4].

§ 5. Transvections.

Let A and C be vectors of L with $AC = 0$ and let $\lambda \in F$. If the mapping

$$X \to X + \lambda(AX)C + \lambda(CX)A, \qquad (X \in L),$$

carries L into L, we shall call it a *double transvection* and denote it by $\mathrm{T}_{A,C,\lambda}$. Since $(\mathrm{T}_{A,C,\lambda}X)(\mathrm{T}_{A,C,\lambda}Y) = XY$, it is an element of $Sp(L)$. Clearly $\mathrm{T}_{A,C,\lambda} = 1_L$ if one of A, C or λ is 0. A double transvection is called *pure of type* (k, l) if it is of the form $\mathrm{T}_{A,C,\lambda}$ with the following properties: (i) there is a canonical basis for L containing both A and C; (ii) A is pure of type k and C is pure of type l; (iii) $A = C$ if $k = l$. Whenever we refer to the fact that $\mathrm{T}_{A,C,\lambda}$ is pure of type (k, l), it will be implicit in this statement that A and C already have these properties unless mentioned otherwise.

The following identities are easy to verify:

$$\mathrm{T}_{A,C,\lambda} = \mathrm{T}_{C,A,\lambda};\ \mathrm{T}_{A,C,\lambda\mu} = \mathrm{T}_{\lambda A,C,\mu} = \mathrm{T}_{A,\lambda C,\mu}; \tag{5.1}$$

$$\mathrm{T}_{A,C,\lambda}\mathrm{T}_{A,C,\mu} = \mathrm{T}_{A,C,\lambda+\mu};\ \mathrm{T}^n{}_{A,C,\lambda} = \mathrm{T}_{A,C,n\lambda} \quad (n \in \mathbf{Z}); \tag{5.2}$$

$$\mathrm{T}_{A,C,\lambda}\mathrm{T}_{A',C',\lambda'} = \mathrm{T}_{A',C',\lambda'}\mathrm{T}_{A,C,\lambda} \text{ if } A,\ C,\ A',\ C' \text{ are mutually orthogonal}; \tag{5.3}$$

$$\sigma\mathrm{T}_{A,C,\lambda}\sigma^{-1} = \mathrm{T}_{\sigma A,\sigma C,\lambda} \text{ if } \sigma \in Sp(L). \tag{5.4}$$

The double transvection $\mathrm{T}_{A,A,\lambda}$ will usually be written $\tau_{A,2\lambda}$ and will be called a *transvection*. Thus $\tau_{A,\lambda}X = X + \lambda(AX)A$. The above identities apply to transvections *mutatis mutandis*. A *pure* transvection is one which is pure as a double transvection. The reader can easily check the following:

$$\mathrm{T}_{A,C,\lambda}\mathrm{T}_{A',C,\lambda} = \tau_{C,\lambda^2(AA')}\mathrm{T}_{A+A',C,\lambda}; \tag{5.5}$$

$$\mathrm{T}_{A,C,\lambda}\mathrm{T}_{A',C,\lambda} = \mathrm{T}_{A+A',C,\lambda} \text{ if } AA' = 0; \tag{5.6}$$

$$\mathrm{T}_{A,C,\lambda} = \tau_{A,\lambda}{}^{-1}\tau_{C,\lambda}{}^{-1}\tau_{A+C,\lambda}; \tag{5.7}$$

$$\tau_{C,\lambda} = \tau_{A,\lambda}{}^{2}\tau_{A+C,\lambda}{}^{2}\tau_{2A+C,\lambda}{}^{-1} \text{ if } AC = 0. \tag{5.8}$$

LEMMA 5.1. *Let* $\mathrm{T} = \mathrm{T}_{A,C,\lambda}$ *be a pure double transvection of type* (k,l). *Then for* $i = 1,\cdots,t$, $o_i(\mathrm{T}) = \mathfrak{I}((\mathrm{T}-1_L)L^{\mathfrak{s}_i})$, *i.e.* $\mathrm{T} \in SSp(L;o(\mathrm{T}))$. *Also* $o_{kl}(\mathrm{T}) = \lambda\mathfrak{s}_k$, *and if* $\{\mathfrak{g}_{ij}\}$ *is a tableau, then* $o(\mathrm{T}) \subseteq \{\mathfrak{g}_{ij}\}$ *if and only if* $o_{kl}(\mathrm{T}) = \lambda\mathfrak{s}_k \subseteq \mathfrak{g}_{kl}$.

Remark. Very often in the rest of the paper we shall introduce a double transvection by means of the notation $\mathrm{T}_{A,C,\lambda}$, and it will then have to be verified that it is indeed a mapping of L into L. To do this it clearly suffices to show that

$$\lambda(AL)C + \lambda(CL)A \subseteq L, \tag{5.9}$$

and almost always this verification will be left (implicitly) to the reader. Whenever the last part of the above lemma is used, however, the fact that $\mathrm{T}L \subseteq L$ follows automatically. Indeed (5.9) is satisfied if and only if both $\lambda\mathfrak{s}_k$ and $\lambda\mathfrak{s}_l$ are in $\mathfrak{o}$, which is certainly the case if $\lambda\mathfrak{s}_k \subseteq \mathfrak{g}_{kl}$.

Proof of Lemma 5.1. Given i, $1 \leqq i \leqq t$, choose a pure vector B of type i which is orthogonal to A and C. Then $\mathrm{T}B = B$ and it follows that $\mathfrak{I}((\mathrm{T}+1_L)L^{\mathfrak{s}_i}) = L^{\mathfrak{s}_i}$ by Lemma 2.1. The first part of the lemma follows from this and (3.3).

Now it is easy to see that $(\mathrm{T}-1_L)L^{\mathfrak{s}_i} = \lambda(AL^{\mathfrak{s}_i})C + \lambda(CL^{\mathfrak{s}_i})A$. Hence from the first part of the lemma and from Lemma 2.3 we get $o_{kl}(\mathrm{T}) = \lambda\mathfrak{s}_k$. Also $o_i(\mathrm{T}) = \lambda(\mathfrak{s}_i \cap \mathfrak{s}_k)L^{\mathfrak{s}_l} + \lambda(\mathfrak{s}_i \cap \mathfrak{s}_l)L^{\mathfrak{s}_k}$. Suppose $\lambda\mathfrak{s}_k \subseteq \mathfrak{g}_{kl}$ and put $\mathfrak{L}_i = (\mathfrak{g}_{i1},\cdots,\mathfrak{g}_{it})$ for $i = 1,\cdots,t$. We may suppose that $k \leqq l$ since $\lambda\mathfrak{s}_k \subseteq \mathfrak{g}_{kl}$ if and only if $\lambda\mathfrak{s}_l \subseteq \mathfrak{g}_{lk}$ by (3.11). If $k \leqq i \leqq l$, then

$$o_i(\mathrm{T}) \subseteq \mathfrak{s}_i\mathfrak{s}_k{}^{-1}\mathfrak{g}_{kl}L^{\mathfrak{s}_l} + \mathfrak{g}_{lk}L^{\mathfrak{s}_k} \subseteq \mathfrak{s}_i\mathfrak{s}_k{}^{-1}\mathfrak{L}_k + \mathfrak{L}_l \subseteq \mathfrak{L}_i.$$

Similarly $o_i(\mathrm{T}) \subseteq \mathfrak{L}_i$ if $i \leqq k$ or $i \geqq l$. This finishes the proof.

§ 6. Extension of Symplectic Mappings from a Component.

Suppose that $L = J \perp K$. We define a monomorphism $\iota\colon Sp(J) \to Sp(L)$ by putting $\iota\sigma = \sigma \perp 1_K$. Clearly double transvections go into double transvections, in particular pure double transvections are transformed into pure double transvections since a canonical basis for J can be extended to one for L. Note also that the normal subgroup of $Sp(L)$ generated by $\iota\sigma$ contains the image under ι of the normal subgroup of $Sp(J)$ generated by σ.

Now let $\{\mathfrak{g}_{ij}\}$ be a tableau for L. If $J \doteq J_1 \perp \cdots \perp J_s$ is a canonical splitting, define I to be the set of integers i, $1 \leqq i \leqq t$, such that $\mathfrak{s}_i = J_k J_k$ for some k. Select those ideals $\mathfrak{g}_{ij}$ for which i and j are in I; they form an s by s array which is easily seen to be admissible as a tableau for J (with the obvious indexing). We shall denote this tableau by $\{\mathfrak{g}_{ij}\}'$. Now re-index the canonical splitting for J in such a way that $J = \mathop{\perp}\limits_{1 \leqq i \leqq t} J_i$ with $J_i J_i = \mathfrak{s}_i$ if $i \in I$, otherwise $J_i = 0$. Define $\mathfrak{I}_i = \sum_{j=1}^{t} \mathfrak{g}_{ij} J_j$ and $\mathfrak{L}_i = \sum_{j=1}^{t} \mathfrak{g}_{ij} L_j$ for $i = 1, \cdots, t$.

Suppose now that $\sigma \in SSp(J; \{\mathfrak{g}_{ij}\}')$. Then $(\sigma - 1_J) J_i \subseteq \mathfrak{I}_i$ for $i = 1, \cdots, t$. So

$$(\iota\sigma - 1_L) L^{\mathfrak{s}_i} = (\sigma - 1_J) J^{\mathfrak{s}_i} \subseteq \sum_{j=1}^{t} ((\mathfrak{s}_i \mathfrak{s}_j^{-1}) \cap \mathfrak{o}) \mathfrak{I}_j$$

by (1.2). But $\mathfrak{I}_j \subseteq \mathfrak{L}_j$ for all j (apply Lemma 2.3 to $\mathbf{X} = \mathfrak{I}_j$). Thus $(\iota\sigma - 1) L^{\mathfrak{s}_i} \subseteq \mathfrak{L}_i$ by (3.6) and (3.7), and we have therefore proved that

$$\iota \colon SSp(J; \{\mathfrak{g}_{ij}\}') \to SSp(L; \{\mathfrak{g}_{ij}\}).$$

§ 7. Generation of *SSp* by Pure Double Transvections.

LEMMA 7.1. *Let B and B' be pure vectors of type i, let $\{\mathfrak{g}_{ij}\}$ be a tableau, and suppose that $B' - B \in (\mathfrak{g}_{i1}, \cdots, \mathfrak{g}_{it})$. Then there exists a product of pure double transvections, each in $SSp(L; \{\mathfrak{g}_{ij}\})$, which carries B to B'. Moreover the types of the pure double transvections involved are all of the form (i, k), $1 \leqq k \leqq t$.*

Proof. 1) Suppose $\mathfrak{g}_{ii} = \mathfrak{o}$. If $[BB'] = \mathfrak{s}_i$, then

$$\tau X = X + (B' - B)((B' - B)X)(B'B)^{-1} \tag{7.1}$$

carries B to B' and is clearly in $Sp(L)$, and $\tau \in SSp = SSp(L; \{\mathfrak{g}_{ij}\})$ by Lemma 5.1. Since $[(B' - B)B] = \mathfrak{s}_i$, $B' - B$ is pure and so τ itself is the required product. Now suppose $[BB'] \subset \mathfrak{s}_i$. Choose a pure vector A of type i with $[AB] = [AB'] = \mathfrak{s}_i$. The pure transvection $\tau = \tau_{A-B,\lambda}$ with $\lambda = (AB)^{-1}$ is in SSp by Lemma 5.1, and $\tau B = A$. Similarly we can find $\tau' \in SSp$ such that $\tau' A = B'$, and then $\tau'\tau$ is the required product.

2) Suppose $\mathfrak{g}_{ii} \subset \mathfrak{o}$. Choose a canonical splitting $L = L_1 \perp \cdots \perp L_t$ with $B \in L_i$, and write $B' - B = \sum_{j=1}^{t} \gamma_j E_j$ with E_j maximal in L_j and $\gamma_j \in \mathfrak{g}_{ij}$

for $j = 1, \cdots, t$. We shall first find a product ρ of the required kind which carries B to $B + \gamma_i E_i$ and is identity on L_j, $j \neq i$. If $[BE_i] = \mathfrak{s}_i$, ρ can be chosen as the transvection

$$\rho X = X + \gamma_i E_i (E_i X)(E_i B)^{-1}.$$

Assume then that $[BE_i] \subset \mathfrak{s}_i$. Choose $A \in L_i$ with $[AB] = \mathfrak{s}_i$. The required product is $\rho = \tau' \tau$ where

$$\begin{aligned} \tau X &= X + \gamma_i (E_i + A)((E_i + A)X)((E_i + A)B)^{-1} \\ \tau' X &= X - \gamma_i A(AX)(A(B + \gamma_i E_i))^{-1}. \end{aligned}$$

(The assumption that $\mathfrak{g}_{ii} \subset \mathfrak{o}$ is used here to show that $[A(B + \gamma_i E_i)] = \mathfrak{s}_i$).

Next suppose that $j \neq i$. We shall find a double transvection T_j of the required kind which carries B to $B + \gamma_j E_j$ and leaves L_k, $k \neq i$ or j, fixed. Namely, with A as above, define

$$\mathrm{T}_j X = X + \gamma_j (AB)^{-1}((AX)E_j + (E_j X)A).$$

We have $\mathrm{T}_j \in SSp$ by Lemma 5.1. The product

$$\rho \prod_{j \neq i} \mathrm{T}_j \tag{7.2}$$

satisfies the conditions of the lemma.

LEMMA 7.2. *In addition to the hypotheses of Lemma 7.1, suppose that A is pure of type i with $[AB] = \mathfrak{s}_i$ and $AB = AB'$. Then the product described in Lemma 7.1 can be chosen to have the additional property that it leaves A fixed.*

Proof. 1) Suppose $\mathfrak{g}_{ii} = \mathfrak{o}$. If $[BB'] = \mathfrak{s}_i$, the transvection (7.1) is the required product. Assume $[BB'] \subset \mathfrak{s}_i$. Define

$$\tau' X = X + (B' - A - B)((B' - A - B)X)(B'(A + B))^{-1}.$$

Then $\tau' \tau_{A,\lambda}$ with $\lambda = (AB)^{-1}$ is the required product.

2) Suppose $\mathfrak{g}_{ii} \subset \mathfrak{o}$. Write $B' - B$ as in 2) of Lemma 7.1. Now $\mathfrak{o}A + \mathfrak{o}B$ splits L by Lemma 1.2 and we may therefore suppose that both A and B are in L_i. Thus $E_j A = 0$ if $j \neq i$, and $\gamma_i E_i A = 0$ since $AB = AB'$. The procedure in 2) of Lemma 7.1 can be applied here almost word for word, once the A in that lemma has been replaced by the A here. The required product is again given by (7.2).

THEOREM 7.3. *If $\{\mathfrak{g}_{ij}\}$ is a tableau, the group $SSp(L; \{\mathfrak{g}_{ij}\})$ is generated by the pure double transvections contained in it.*

Remarks. This theorem is false if "pure double transvections" is replaced by "transvections," whence the necessity of introducing double transvections. As an example consider a lattice L with canonical splitting $L = L_1 \perp L_2$ and $\mathfrak{s}_1 = \mathfrak{o}$, $\mathfrak{s}_2 = \mathfrak{p}$. Consider the tableau $\mathfrak{g}_{11} = \mathfrak{g}_{21} = \mathfrak{g}_{22} = \mathfrak{p}$ and $\mathfrak{g}_{12} = \mathfrak{o}$. Using Lemma 5.1, one can easily show that $o_{12}(\tau) \subseteq \mathfrak{p}$ for all transvections $\tau \in SSp$; it follows by Lemma 3.2 that the group G generated by the transvections in SSp has $o_{12}(G) \subseteq \mathfrak{p}$. But SSp contains a mapping σ with $o_{12}(\sigma) = \mathfrak{o}$, namely the pure double transvection $\mathrm{T}_{A,C,1}$ where A is maximal in L_1 and C is maximal in L_2. Thus $G \neq SSp$.

There are however some particular tableaus whose special congruence subgroups are generated by transvections, or even pure transvections. Indeed one can show that $SSp(L; \{\mathfrak{g}_{ij}\})$ is generated by the pure transvections in it (resp. the transvections in it) if and only if

$$\mathfrak{g}_{ij} \subseteq \mathfrak{g}_{ii} + \mathfrak{g}_{i+1,i+1} + \cdots + \mathfrak{g}_{jj} \text{ for all } i \text{ and } j, \qquad 1 \leqq i \leqq j \leqq t.$$

(resp. $\mathfrak{g}_{ij}\mathfrak{g}_{ji} \subseteq \mathfrak{g}_{ii}\mathfrak{g}_{jj}$ for all i and j). The sufficiency is most easily proved by expressing a pure double transvection in SSp as a product of pure transvections (resp. transvections) in SSp and then applying Theorem 7.3. These results are not necessary for the rest of the paper, and proofs are therefore omitted. We shall merely deal with a special case of them in Corollary 7.4.

Proof of Theorem 7.3. Let $\sigma \in SSp$ and suppose that

$$0 \neq L = (\mathfrak{o}A + \mathfrak{o}B) \perp J \text{ with } [AB] = \mathfrak{s}_1.$$

By the definition of SSp, $\sigma A - A$ and $\sigma B - B$ are in $\mathfrak{L}_1 = (\mathfrak{g}_{11}, \cdots, \mathfrak{g}_{1t})$. Thus by Lemma 7.1 there is a product Π_1 of pure double transvections, each in SSp, which carries A to σA—and B to B', say. Therefore $B' - B \in \mathfrak{L}_1$ and so $B' - \sigma B = (B' - B) - (\sigma B - B)$ is also in $\mathfrak{L}_1$. Since $(\sigma A)B' = AB = \sigma A \sigma B$, we can find another such product Π_2 by Lemma 7.2, which carries B' to σB and leaves σA fixed. Then $\Pi = \Pi_2\Pi_1$ carries A to σA and B to σB. Thus $\phi = \Pi^{-1}\sigma \in SSp$ and is identity on $\mathfrak{o}A + \mathfrak{o}B$ and hence $\phi' = \phi \mid J \in Sp(J)$. If $J = 0$, the theorem is proved, so assume that $\dim L > 2$, and that the theorem holds for lattices with dimension $< \dim L$.

We now apply the procedure of § 6 to the present situation; in particular $\{\mathfrak{g}_{ij}\}'$ will stand for the tableau for J "induced" by $\{\mathfrak{g}_{ij}\}$. If $1 \leqq k$, $l \leqq t$, then $((\phi - 1_L)L^{\mathfrak{s}_k})L^{\mathfrak{s}_l} = ((\phi' - 1_J)J^{\mathfrak{s}_k})J^{\mathfrak{s}_l}$, and it follows easily (by Lemma 2.3) that $\phi' \in SSp' = SSp(J; \{\mathfrak{g}_{ij}\}')$. By the induction hypothesis ϕ' is a product of pure double transvections, each in SSp', whence $\phi = \iota\phi'$ and $\sigma = \Pi\phi$ are products of pure double transvections, each of which is in SSp, as required.

COROLLARY 7.4. *$Sp(L)$ is generated by pure transvections.*

Proof. Apply (5.7) and Lemma 5.1 to show that every pure double transvection is a product of pure transvections. The corollary then follows from the theorem.

§ 8. Determination of the Normal Subgroups of $Sp(L)$.

From now on both assumptions, char $\mathfrak{o}/\mathfrak{p} \neq 2$ and $\mathfrak{o}/\mathfrak{p} \neq \boldsymbol{F}_3$, are in force. If $\sigma \in Sp(L)$, we define $G(\sigma)$ to be the normal subgroup of $Sp(L)$ generated by σ.

LEMMA 8.1. *Let T and T' be pure double transvections, both of type (k, l). Then $\mathrm{T}' \in G(\mathrm{T})$ if $o(\mathrm{T}') \subseteq o(\mathrm{T})$.*

Proof. Suppose $\mathrm{T} = \mathrm{T}_{A,C,\lambda}$, $\mathrm{T}' = \mathrm{T}_{A',C',\lambda'}$. Pick a canonical splitting with $A \in L_k$ and $C \in L_l$. Let us first deal with $k \neq l$. If $A = A'$, $C = C'$ and $\lambda' = \epsilon\lambda$ for some unit ϵ, we may choose $\sigma \in Sp(L)$ with $\sigma A = A'$ and $\sigma C = \epsilon C'$ (since $\{A, C\}$ and $\{A', \epsilon C'\}$ are in canonical bases), whence $\mathrm{T}' = \sigma\mathrm{T}\sigma^{-1} \in G(\mathrm{T})$ by (5.4) and (5.1). If $A = A'$, $C = C'$ and $\lambda' = \alpha\lambda$ with $\alpha \in \mathfrak{p}$, then $\mathrm{T}' = \mathrm{T}\mathrm{T}_{A,C,(\alpha-1)\lambda} \in G(\mathrm{T})$ by (5.2) since $\alpha - 1$ is a unit. The general case now follows easily. Namely choose $\rho \in Sp(L)$ with $\rho A = A'$, $\rho C = C'$; by (5.4), $\mathrm{T}' = \rho\mathrm{T}_{A,C,\lambda'}\rho^{-1} \in G(\mathrm{T})$.

Now suppose $k = l$, so that $A = C$, $A' = C'$, and T and T' are pure transvections. Clearly $\mathrm{T} \mid L_k \in Sp(L_k)$ and so $G(\mathrm{T} \mid L_k)$ $(\subseteq Sp(L_k))$ contains all transvections of the form $\tau_{A,\mu} \in Sp(L_k)$ with $\mu \in [\lambda]$, by Proposition 3.1 in [6]. By § 6, $G(\mathrm{T})$ contains all transvections $\tau_{A,\mu} \in Sp(L)$ with $\mu \in [\lambda]$. Choose $\rho \in Sp(L)$ with $\rho A = A'$. Then $\mathrm{T}' = \rho\tau_{A,2\lambda'}\rho^{-1} \in G(\mathrm{T})$ as required.

LEMMA 8.2. *Suppose that $\sigma \in Sp(L)$, and that $\sigma A = \alpha A + \beta B$ where $\mathfrak{o}A + \mathfrak{o}B$ splits L and $[AB] = \mathfrak{s}_i$. Then $G(\sigma)$ contains all pure transvections τ of type (i, i) with $o_{ii}(\tau) \subseteq [\beta]$.*

Proof. Put $\mu = (AB)^{-1}$, $\tau = \tau_{A,\mu}$, and $\rho = \sigma\tau\sigma^{-1}\tau^{-1} \in G(\sigma)$. Now $\rho = \tau_{\alpha A+\beta B,\mu}\tau^{-1}$ by (5.4) whence $\rho(\mathfrak{o}A + \mathfrak{o}B) = \mathfrak{o}A + \mathfrak{o}B$ and $\rho \mid (\mathfrak{o}A + \mathfrak{o}B)^{0}$ is identity. A direct calculation shows that $o(\rho \mid \mathfrak{o}A + \mathfrak{o}B) \supseteq \{[\beta]\}$. The lemma follows by Proposition 3.1 in [6], and Lemma 8.1.

LEMMA 8.3. *Let $\rho \in Sp(L)$ and suppose that $o(\rho) = \{\mathfrak{g}_{ij}\}$ with $\mathfrak{g}_{ii} \subset \mathfrak{o}$. Then there exists a pure vector E of type i in L and $\sigma \in G(\rho)$ such that $[(\sigma E)E] = \mathfrak{s}_i\mathfrak{g}_{ii}$.*

Proof. By Lemma 2.3 there exists $A \in L^{\mathfrak{s}_i}$ such that $((\rho - \delta_L)A)L^{\mathfrak{s}_i}$

$= \mathfrak{s}_i\mathfrak{g}_{ii}$, where $\Delta_i(\rho) = \delta$. Splitting A into components with respect to a canonical splitting, we see that we may suppose that A is pure of type i: let $[AB] = \mathfrak{s}_i$ with $B \in L^{\mathfrak{s}_i}$. Since $(\rho A)A = ((\rho - \delta_L)A)A$, we see that $(\rho A)A \in \mathfrak{s}_i\mathfrak{g}_{ii}$. If $[(\rho A)A] = \mathfrak{s}_i\mathfrak{g}_{ii}$, we are finished, so assume

$$[(\rho A)A] \subset \mathfrak{s}_i\mathfrak{g}_{ii}. \tag{8.1}$$

Let $C \in L^{\mathfrak{s}_i}$ satisfy $[((\rho - \delta_L)A)C] = \mathfrak{s}_i\mathfrak{g}_{ii}$ and define E as follows: put $E = B$ if $[((\rho - \delta_L)A)B] = \mathfrak{s}_i\mathfrak{g}_{ii}$; otherwise put $E = C$ if $[AC] = \mathfrak{s}_i$ or $E = B + C$ if $[AC] \subset \mathfrak{s}_i$. Therefore $[((\rho - \delta_L)A)E] = \mathfrak{s}_i\mathfrak{g}_{ii}$, $E \in L^{\mathfrak{s}_i}$, and $[AE] = \mathfrak{s}_i$, whence

$$(\rho A)E = (AE)(\delta + \gamma), [\gamma] = \mathfrak{g}_{ii}. \tag{8.2}$$

Now put $\lambda = (AE)^{-1}$ and define

$$\sigma = \rho\tau_{A,\lambda}\rho^{-1}\tau^{-1}_{A,\lambda} = \tau_{\rho A,\lambda}\tau^{-1}_{A,\lambda} \in G(\rho).$$

A direct calculation using (8.1) and (8.2) shows that $[(\sigma E)E] = \mathfrak{s}_i\mathfrak{g}_{ii}$ as required.

LEMMA 8.4. *Suppose that* $\rho \in Sp(L)$. *Then* $G(\rho)$ *contains all pure transvections* τ *of type* (i,i) *with* $\mathfrak{o}_{ii}(\tau) \subseteq \mathfrak{o}_{ii}(\rho)$.

Proof. The proof is very similar to that of Lemma 3.2 in [6], and we therefore only sketch it. Note that (i) of that lemma corresponds to our Lemma 8.2.

Suppose that $\mathfrak{o}_{ii}(\rho) = \mathfrak{o}$. We then determine $A \in L^{\mathfrak{s}_i}$ such that $AL^{\mathfrak{s}_i} = ((\rho - 1_L)A)L^{\mathfrak{s}_i} = \mathfrak{s}_i$, and such that A and $(\rho - 1_L)A$ are independent $\operatorname{mod}(L^{\mathfrak{s}_i})^{\mathfrak{p}\mathfrak{s}_i}$—that is to say, the images of A and $(\rho - 1_L)A$ in the vector space $L^{\mathfrak{s}_i}/(L^{\mathfrak{s}_i})^{\mathfrak{p}\mathfrak{s}_i}$ over $\mathfrak{o}/\mathfrak{p}$ are independent. The rest follows as in (ii) of the aforementioned lemma.

Now consider $\mathfrak{o}_{ii}(\rho) \subset \mathfrak{o}$. By Lemma 8.3 we can find a pure vector A of type i and $\sigma \in G(\rho)$ such that $[(\sigma A)A] = \mathfrak{s}_i\mathfrak{o}_{ii}(\rho)$. Choose $B \in L^{\mathfrak{s}_i}$ with $[AB] = \mathfrak{s}_i$; let $\lambda = (AB)^{-1}$ and put $\tau = \tau_{A,\lambda}$. We then define $\sigma' = \sigma^{-1}\tau\sigma\tau^{-1}$. The rest of the proof follows as in (iii) in Klingenberg's lemma, and we therefore omit it.

LEMMA 8.5. *If* $\sigma \in Sp(L)$, *then the normal subgroup* $G(\sigma)$ *generated by* σ *contains* $SSp(L;\mathfrak{o}(\sigma))$.

Proof. By Theorem 7.3, it suffices to show that $G(\sigma)$ contains all pure double transvections with order $\subseteq \mathfrak{o}(\sigma) = \{\mathfrak{g}_{ij}\}$, say. Order the $\frac{1}{2}t(t+1)$ pairs (i,j), $1 \leqq i \leqq j \leqq t$, lexicographically. The proof is by induction on the type (i,j) of the pure double transvections. By Lemma 8.4, $G(\sigma)$ contains all pure transvections of type (i,i) in SSp. Therefore suppose that $(i,j) > (1,1)$, that $i < j$, and that $G(\sigma)$ contains all pure double transvections in SSp whose type (i',j') satisfies $(i',j') < (i,j)$.

During the proof we shall use the mapping $\Delta_m \colon GSp \to \{\pm 1\}$ (introduced in the proof of Theorem 4.1) which is defined for all m such that $\mathfrak{g}_{mm} \subset \mathfrak{o}$.

1) Define k to be the largest integer $\leqq t$ such that $\mathfrak{s}_k \mathfrak{g}_{ik} = \mathfrak{s}_j \mathfrak{g}_{ij}$; thus

$$\mathfrak{s}_l \mathfrak{g}_{il} \subset \mathfrak{s}_k \mathfrak{g}_{ik}, \quad k < l \leqq t. \tag{8.3}$$

Suppose that $j < k$ and that $\mathrm{T}_{A,B,\lambda}$ is a pure double transvection in SSp of type (i,j). Extend $\{A,B\}$ to a canonical basis and let C and D be members of this basis of type k with $[CD] = \mathfrak{s}_k$. If $\beta \mathfrak{s}_j = \mathfrak{s}_k$, then $\mathfrak{o}(C + \beta B) + \mathfrak{o}D$ splits $L_j \perp L_k$ by Lemma 1.2 where (1.1) is the canonical splitting arising from this basis. Thus A and $C + \beta B$ are both in some canonical basis of L. Now

$$\mathrm{T}_{A,B,\lambda} = \mathrm{T}_{A,C+\beta B,\mu} \mathrm{T}_{A,C,\mu}{}^{-1}, \quad \mu = \beta^{-1}\lambda.$$

The definition of k and Lemma 5.1 show that the double transvections on the right side are in SSp if $\mathrm{T}_{A,B,\lambda}$ is in SSp. Therefore it suffices to show that $G(\sigma)$ contains all pure double transvections in SSp of type (i,k), and this is what we shall do.

2) Suppose that $\mathfrak{g}_{il} = \mathfrak{g}_{ik}$ for some l, $i \leqq l < k$ (so that, in particular, $j = k$). Let $\mathrm{T}_{A,C,\lambda} \in SSp$ be pure of type (i,k). Choose a basis of L which contains A and C, and let B be one of its members which is pure of type l; in particular choose $B = A$ if $l = i$. Now $\mathrm{T}_{A,C,\lambda} = \mathrm{T}_{A,B+C,\lambda} \mathrm{T}_{A,B,\lambda}{}^{-1}$ and it follows that $\mathrm{T}_{A,C,\lambda} \in G(\sigma)$ by the induction hypothesis if $l > i$. Suppose $l = i$. Then $\mathrm{T}_{A,A,\lambda} = \tau_{A,2\lambda}$ and $\mathrm{T}_{A,A+C,\lambda} = \tau_{2A+C,\lambda} \tau_{A,\lambda}{}^{-1} \tau_{A+C,\lambda}{}^{-1}$ by (5.7), and so we again have $\mathrm{T}_{A,C,\lambda} \in G(\sigma)$ by induction. Therefore from now on we may assume that

$$\mathfrak{g}_{il} \subset \mathfrak{g}_{ik} \text{ if } l < k. \tag{8.4}$$

3) Suppose $\mathfrak{g}_{kk} = \mathfrak{g}_{ik}$. Let $\mathrm{T} = \mathrm{T}_{A,C,\lambda}$ be a pure double transvection in SSp of type (i,k). If $\alpha \mathfrak{s}_i = \mathfrak{s}_k$, $\mathrm{T} = \mathrm{T}_{C+\alpha A,C,\mu} \mathrm{T}_{C,C,\mu}{}^{-1}$ where $\mu = \alpha^{-1}\lambda$. Write $\mathrm{T}_{C+\alpha A,C,\mu}$ as a product of pure transvections of type (k,k), using (5.7). We know already that all such pure transvections which are in SSp are also in $G(\sigma)$, and it follows that $\mathrm{T} \in G(\sigma)$, again by several applications of Lemma 5.1. So from now on we may also assume that

$$\mathfrak{g}_{kk} \subset \mathfrak{g}_{ik} \subseteq \mathfrak{o} \tag{8.5}$$

4) If $\Delta_i(\sigma) = \delta$, $\delta = \pm 1$, we can find a pure vector A of type i such that $((\sigma - \delta_L)A)L^{\mathfrak{s}_k} = \mathfrak{s}_k \mathfrak{g}_{ik}$. Choose a canonical splitting $L = L_1 \perp \cdots \perp L_t$ so that $A \in L_i$. Write $(\sigma - \delta_L)A = \sum_{l=1}^{t} \gamma_l E_l$ where E_l is maximal in L_l and $\gamma_l \in \mathfrak{g}_{il}$, for $l = 1, \cdots, t$. It follows from (8.3) and (8.4) that $((\sigma - \delta_L)A)L^{\mathfrak{s}_k} = \gamma_k \mathfrak{s}_k$, whence

$$[\gamma_k] = \mathfrak{g}_{ik}. \tag{8.6}$$

Put $\gamma_l = \alpha_l \gamma_k$ for $j \leqq l < k$. By definition of k, $\alpha_l \in \mathfrak{s}_k \mathfrak{s}_l^{-1}$, whence $\sum_{l=j}^{k-1} \alpha_l E_l \in L^{\mathfrak{s}_k}$. Thus choosing a new canonical splitting for $L_j \perp \cdots \perp L_k$ if necessary (and hence also for L), we may suppose that $\gamma_l = 0$ for $j \leqq l < k$.

We now apply the technique in § 6 with $J = L_1 \perp \cdots \perp L_{j-1}$, and with the tableau $\{\mathfrak{g}_{ij}\}$ as we have defined it here; $\{\mathfrak{g}_{ij}\}'$ is the "induced" tableau for J. According to Lemma 7.1 there is a product Π' of pure double transvections, each in $SSp(J; \{\mathfrak{g}_{ij}\}')$ and of type (i, l), $1 \leqq l < j$, which carries $\delta A + \sum_{l=1}^{j-1} \gamma_l E_l$ to δA. Therefore $\Pi = \iota \Pi' \in SSp(L; \{\mathfrak{g}_{ij}\})$ is a product of pure double transvections of the same types, which carries $\delta A + \sum_{l=1}^{j} \gamma_l E_l$ to δA, and is identity on J^0. Moreover since these double transvections are all in SSp, they are also in $G(\sigma)$ by the induction hypothesis. Thus $\sigma' = \Pi\sigma \in G(\sigma)$ and $\sigma' A = \delta A + \sum_{l=k}^{t} \gamma_l E_l$. In particular $(\sigma' A)A = 0$.

5) Choose $B \in L_i$ so that $\mathfrak{o}A + \mathfrak{o}B$ splits L_i. Note that $(\sigma' A)B = \delta AB$. Put $\alpha = (AB)^{-1}$ and $\tau = \tau_{A,\delta\alpha}$. Then

$$\rho = \sigma' \tau \sigma'^{-1} \tau^{-1} \in G(\sigma), \quad \rho B = B + (\sigma' - \delta_L)A, \text{ and } \rho A = A.$$

Since $\mathfrak{g}_{kk} \subset \mathfrak{o}$ and $\mathfrak{g}_{ii} \subset \mathfrak{o}$, we have

$$\Delta_i(\rho) = \Delta_k(\rho) = 1 \tag{8.7}$$

by Lemma 3.2. Notice too, that $(\rho B)B = 0$ and $(\rho B)A = BA$.

For the rest of the proof we shall write $T_C = T_{\rho B, C, \alpha}$ for a (suitable) vector C and $\alpha = (AB)^{-1}$. Now $\rho' = T_B^{-1} \rho T_B \rho^{-1} \in G(\sigma)$ and clearly $\rho' = T_{\rho^2 B - B}$. It is easy to see that

$$\rho^2 B - B = \sum_{l=k}^{t} C_l \text{ where } C_l = \gamma_l(\rho E_l + E_l). \tag{8.8}$$

Using this expression, one can easily check that $(\rho^2 B - B)L^{\mathfrak{s}_k} = \mathfrak{s}_k \mathfrak{g}_{ik}$ by

(8.7), (8.5), and the properties of $\gamma_k, \cdots, \gamma_t$. Also $\Delta_i(\rho^2) = 1$, whence $(\rho^2 B - B)\rho A \in \mathfrak{s}_i \mathfrak{g}_{ii} \subset \mathfrak{s}_i \mathfrak{g}_{ik}$. It follows by a straightforward calculation that $((\rho' - 1_L)\rho A)L^{\mathfrak{s}_k} = \mathfrak{s}_k \mathfrak{g}_{ik}$. Since $\Delta_i(\rho') = 1$, we get $o_{ik}(\rho') = \mathfrak{g}_{ik}$.

6) It is obvious that $(\rho B)C_l = 0$ for all l since $(\rho B)\rho E_l = (\rho B)E_l = 0$, and only slightly harder to show that

$$C_l C_m \in \mathfrak{s}_i \mathfrak{g}_{ii} \text{ for all } l \text{ and } m, \tag{8.9}$$

by using (3.9) and 3.10).

Now let $\Sigma_k = \sum_{l=k+1}^{t} C_l$, and put $\gamma = \alpha^2(\Sigma_k C_k)$. By (5.5), $\rho' = \tau_{\rho B, \gamma}{}^{-1} \rho_k$ where $\rho_k = \mathrm{T}_{\Sigma_k} \mathrm{T}_{C_k}$. From (8.9) and Lemma 5.1 we see that $\tau_{\rho B, \gamma} \in SSp$ —and so is also in $G(\sigma)$—and that $o_{ik}(\tau_{\rho B, \gamma}) \subset \mathfrak{g}_{ik}$, using (8.4) also. Thus $\rho_k \in G(\sigma)$ and $o_{ik}(\rho_k) = \mathfrak{g}_{ik}$. We now split up $\Sigma_k = \Sigma_{k+1} + C_{k+1}$, and proceed in a similar manner. Eventually we find

$$\rho_{t-1} = \mathrm{T}_{C_t} \cdots \mathrm{T}_{C_k} \in G(\sigma), o_{ik}(\rho_{t-1}) = \mathfrak{g}_{ik} \tag{8.10}$$

7) Let us choose a new canonical splitting $L = L_1' \perp \cdots \perp L_t'$ such that $\rho B \in L_i'$. For the moment consider l to be fixed, $l \geqq k$. Then we can write $C_l = \sum_{m=1}^{t} D_m$ with $D_m \in \mathfrak{g}_{im} L_m'$ by (2.4) in Theorem 2.2, (8.8), and the fact that $o(\rho^2) \subseteq \{\mathfrak{g}_{ij}\}$. Clearly $(\rho B)D_m = 0$ for all m. Thus by (5.6) we can write T_{C_l} as the product $\mathrm{T}_{D_1} \cdots \mathrm{T}_{D_t}$. Now factor each double transvection in (8.10) in this manner. Therefore by (5.3) we may suppose that ρ_{t-1} is written as the product of double transvection of the form T_D where the first few factors in the product have $D \in \mathfrak{g}_{i1} L_1'$, the next few have $D \in \mathfrak{g}_{i2} L_2', \cdots,$ the last few have $D \in \mathfrak{g}_{it} L_t'$.

Now take all those T_D's with $D \in \mathfrak{g}_{im} L_m'$ (fixed m) and multiply them together two at a time. One gets, for example,

$$\mathrm{T}_D \mathrm{T}_{D'} = \tau_{\rho B, \lambda} \mathrm{T}_{D+D'}, \ \lambda = \alpha^2(DD'),$$

by (5.5). Since $(\rho B)D = 0$ for any D, $\tau = \tau_{\rho B, \lambda}$ commutes with everything in the product. Moreover one can show, as in 6), that $\tau \in G(\sigma)$ and that $o_{ik}(\tau) \subset \mathfrak{g}_{ik}$. Thus ultimately we can find $\pi \in G(\sigma)$ with $o_{ik}(\pi) = \mathfrak{g}_{ik}$ of the form

$$\pi = \mathrm{T}_{F_1} \cdots \mathrm{T}_{F_t}, F_m \in \mathfrak{g}_{im} L_m' \text{ for } 1 \leqq m \leqq t.$$

A direct calculation of $((\mathrm{T}_{F_m} - 1_L)L^{\mathfrak{s}_i})L^{\mathfrak{s}_k}$, using (8.3) and (8.4), shows $o_{ik}(\mathrm{T}_{F_m}) \subset \mathfrak{g}_{ik}$ for $m \neq k$. Therefore we must have $o_{ik}(\mathrm{T}_{F_k}) = \mathfrak{g}_{ik}$. Now choose $\mu \in Sp(L)$ which is identity on L_m' for $m \neq k$, and which sends F_k onto $2F_k$. Then

$$T_{F_k} = \mu\pi\mu^{-1}\pi^{-1} \in G(\sigma).$$

By Lemma 8.1 this finishes the proof of Lemma 8.5.

Structure Theorem for *Sp*(*L*).

(i) *Let G be a normal subgroup of $Sp(L)$ with order $o(G) = \{\mathfrak{g}_{ij}\}$. Then*

$$SSp(L;\{\mathfrak{g}_{ij}\}) \subseteq G \subseteq GSp(L;\{\mathfrak{g}_{ij}\}). \tag{8.11}$$

(ii) *Conversely any group G which satisfies* (8.11) *is a normal subgroup of $Sp(L)$ of order $\{\mathfrak{g}_{ij}\}$.*

(iii) *If GSp and SSp are congruence groups of the same order, then GSp/SSp is a finite direct product of the cyclic group Z_2 of two elements with itself* (*cf. Theorems* 4.1 *and* 4.2).

(iv) *If G is a normal subgroup of $Sp(L)$ not contained in $\{\pm 1_L\}$, the centralizer of G in $Sp(L)$ is $\{\pm 1_L\}$. Then center of G is $G \cap \{\pm 1_L\}$. The center of any general congruence subgroup is $\{\pm 1_L\}$ while the center of any proper special congruence subgroup is $\{1_L\}$.*

(v) *If G is a normal subgroup of $Sp(L)$, then the mixed commutator subgroup $[Sp(L), G]$ is equal to $SSp(L; o(G))$. In particular $Sp(L)$ is equal to its derived group.*

Proof. (i) Let T be a pure double transvection of type (k,l) in $SSp(L; o(G))$. Because of the definition of order, we can find σ in G such that $o_{kl}(\mathrm{T}) \subseteq o_{kl}(\sigma)$. But then $o(\mathrm{T}) \subseteq o(\sigma)$ by Lemma 5.1 and so $\mathrm{T} \in G(\sigma) \subseteq G$ by Lemma 8.5. Thus (i) follows by Theorem 7.3 (and the definition of GSp).

(ii) and (iii). By Theorem 4.1 it suffices to show that $o(SSp(L;\{\mathfrak{g}_{ij}\})) = \{\mathfrak{g}_{ij}\}$. But this follows from Lemma 5.1 and the fact that given $\gamma \in \mathfrak{g}_{kl}$, there exists a pure double transvection T of type (k,l) with $o_{kl}(\mathrm{T}) = [\gamma]$.

(iv) One shows easily, as in Proposition 2.1 of [6], that any element of $Sp(L)$ of the form

$$X \to \alpha(X)X, \ \alpha(X) \in F, \quad \text{for all } X \in L,$$

must be $\{\pm 1_L\}$. Thus $o(G)$ is not the zero tableau. So given a non-zero vector A of L, we can find $\tau_{A,\lambda} \in G$, $\tau_{A,\lambda} \neq 1_L$, by (i). The first part of (iv) now follows as in Proposition 2.1, [6]. The rest follows easily—note that the only special congruence group containing -1_L is $Sp(L)$ itself since $\mathfrak{I}((-1_L - 1_L)L^{\mathfrak{s}_i}) = L^{\mathfrak{s}_i}$.

(v) If $\sigma, \rho \in Sp(L)$ and $\delta = \pm 1$, $\mathfrak{I}((\sigma - \delta_L)L^{\mathfrak{s}_i}) = \mathfrak{I}((\rho\sigma\rho^{-1} - \delta_L)L^{\mathfrak{s}_i})$ for all i. It follows easily from the definitions and Lemma 3.2 that $[Sp(L), G] \subseteq SSp(L; o(G))$. Conversely given a pure double transvection T in $SSp(L; o(G)) \subseteq G$, it is easy to find $\rho \in Sp(L)$ such that $o(\rho \mathrm{T} \rho^{-1} \mathrm{T}^{-1}) = o(\mathrm{T})$—for example by using (5.4), (5.6), and Lemma 5.1. Then (v) follows immediately by (i).

UNIVERSITY OF NOTRE DAME.

REFERENCES.

[1] E. Artin, *Geometric Algebra*, Interscience, New York, 1957.

[2] L. E. Dickson, "Theory of linear groups in an arbitrary field," *Transactions of the American Mathematical Society*, vol. 2 (1901), pp. 363-394.

[2] J. Dieudonné, *Sur les Groupes Classiques*, Hermann, Paris, 1958.

[4] M. Hall, Jr., *The Theory of Groups*, Macmillan, New York, 1959.

[5] C. Jordan, *Traité des Substitutions et des Équations Algébriques*, Gauthier-Villars, Paris, 1870.

[6] W. Klingenberg, "Symplectic groups over local rings," *American Journal of Mathematics*, vol. 85 (1963), pp. 232-240.

[7] C. R. Riehm, "Symplectic groups over discrete valuation rings," *Bulletin of the American Mathematical Society*, vol. 71 (1965), pp. 388-392.

ON ISOLATED RATIONAL SINGULARITIES OF SURFACES.

By MICHAEL ARTIN.*

In 1934, Du Val [3] listed the configurations of curves which can be obtained by resolving certain isolated double points of embedded surfaces (they are depicted in the figure below). These configurations arise naturally in other contexts, for instance as exceptional curves for pluricanonical embeddings of surfaces [7], and so it seems desirable to have a converse result, showing that a singularity giving rise to such a configuration is necessarily a double point. We have reconsidered the question in a more general context, and obtain in addition the correct numerical characterization of "rational" singularities (cf. definition below). This characterization is made without the assumption that the surface is embedded and points out the connection of Du Val's work with Castelnuovo's criterion for exceptional curves [2], ([8], p. 38). Finally, we list the configurations obtained from rational triple points.

Denote by $\bar{V}$ the spectrum of a normal two-dimensional local ring A with algebraically closed residue field k and by

$$\pi\colon V \to \bar{V}$$

a birational proper map with V regular, i.e., a "resolution" of the singularity of $\bar{V}$. According to results of Lichtenbaum [5] much of the theory of curves on algebraic surfaces can be carried over to regular two-dimensional schemes, and we will apply these results freely to V.

We shall call the singularity of $\bar{V}$ *rational* if the higher direct image $R^1\pi^*\mathcal{O}_V$ is zero. This condition is known to be independent of the resolution π of the singularity (because of the theorem on factorization of a birational proper morphism of regular two-dimensional schemes into locally quadratic transformations [9], [5]). In case A is a local ring of a complete surface $\bar{S}$, the condition means that the arithmetic genus of the surface is not changed by resolving the singularity.

Let v be the closed point of $\bar{V}$, let X be the reduced curve whose support is the inverse image of v on V, and let $\{X_i\}$, $i = 1, \cdots, n$, be the

---

Received October 2, 1964.

[1] This research was supported in part by grant NSF GP-1740 of the National Science Foundation.

irreducible components of X. One has the following criterion, of which we will obtain a better version later:

PROPOSITION 1. *$\bar{V}$ has a rational singularity if and only if for each positive divisorial cycle $Z = \sum r_i X_i$ (r_i integers $\geqq 0$) with support in X, $p(Z) \leqq 0$.*

Here $p(Z) = 1 - \chi(Z)$ is the arithmetic genus (cf. [1]).

Proof. The coherent sheaf $R^1\pi^*\mathcal{O}_V$ on $\bar{V}$ is concentrated at the point v. Applying Grothendieck's theorem on holomorphic functions ([4], (4.2.1)) one finds that its stalk at v is

$$\underset{(r)\to(\infty)}{\underleftarrow{\lim}}\; H^1(X, \mathcal{O}_{Z_{(r)}})$$

where $(r) = (r_1, \cdots, r_n)$ and $Z_{(r)} = \sum r_i X_i$. Since X has dimension 1, the map

$$H^1(X, \mathcal{O}_{Z_{(r)}}) \to H^1(X, \mathcal{O}_{Z_{(s)}}),$$

defined whenever $r_i \geqq s_i$ for all i, is surjective. Therefore the singularity is rational if and only if $H^1(X, \mathcal{O}_Z) = 0$ for all positive divisorial cycles $Z = \sum r_i X_i$. The proposition now follows from ([1], Theorem (1.7)) (in that paper we assumed that the curve X was on an algebraic surface, but the proof works in the general case using Lichtenbaum's theory [5]).

Recall that according to Mumford [6] the intersection matrix $\|(X_i \cdot X_j)\|$ is negative definite. The following equivalent condition is useful. (A large part of the proposition is included in [10], Lemmas 7.1 and 7.5. See also [6], [3].):

PROPOSITION 2. *Let $\{X_i\}$ $i = 1, \cdots, n$ be a connected bunch of complete curves on a regular two-dimensional scheme:*

(i) *Suppose that $\|(X_i \cdot X_j)\|$ is negative definite. Then there exist positive cycles $Z = \sum r_i X_i$ which have the property that $(Z \cdot X_i) \leqq 0$ for all i.*

(ii) *Conversely, if there exists a positive cycle $Z = \sum r_i X_i$ such that $(Z \cdot X_i) \leqq 0$ for all i, then $\|(X_i \cdot X_j)\|$ is negative semi-definite, and if in addition $(Z^2) < 0$, then $\|(X_i \cdot X_j)\|$ is negative definite.*

Proof. (i) Suppose that $\|(X_i \cdot X_j)\|$ is negative definite. Then it is clear that there exist cycles Z such that $(Z \cdot X_i) < 0$ for all i. Because of negativeness, $Z > 0$. In fact, if $Z = A - B$ where A, B are non-negative and without common components, one has $(A \cdot B) \geqq 0$, $(B^2) \leqq 0$, hence $(A - B \cdot B) \geqq 0$. Therefore all the inequalities are equalities, and so $B = 0$.

(ii) Suppose that a positive cycle Z exists with $(Z \cdot X_i) \leqq 0$ for all i (resp. $(Z \cdot X_i) \leqq 0$ for all i and $(Z^2) < 0$). We shall prove the assertion by induction on the number n of components of X. If there exists a Y with

$(Y^2) > 0$ (resp. $(Y^2) \geqq 0$ and $Y \neq 0$) then there exists such a Y which is positive. (In fact, if $Y = A - B$ where A, B are non-negative and without common components, then either A or B is the desired cycle.) Moreover it is clear that one may choose Y with connected support, say

$$\operatorname{supp}(Y) = X' = X_1 \cup \cdots \cup X_\nu, \quad \nu \leqq n.$$

Write

$$Z = \sum_{i=1}^{n} r_i X_i$$

and put

$$Z' = \sum_{i=1}^{\nu} r_i X_i.$$

Then clearly $(Z' \cdot X_i) \leqq 0$ for $i = 1, \cdots, \nu$. Suppose we want only to prove that the intersection matrix is negative semi-definite. Then the induction hypothesis contradicts the existence of such a Y with $(Y^2) > 0$ if $\nu < n$, and so we conclude $\operatorname{supp}(Y) = X$. Choose positive integral multiples aY, bZ such that $aY \geqq bZ$ but $\operatorname{supp}(aY - bZ) < X$. We have

$$(aY - bZ)^2 = a^2(Y^2) - bZ(aY - bZ) \geqq a^2 Y^2.$$

Hence $(Y^2) > 0$ implies $(aY - bZ)^2 > 0$ which is a contradiction since $\operatorname{supp}(aY - bZ) < X$.

Now suppose we want to show negative definiteness. We already know that $\|(X_i \cdot X_j)\|$ is negative semi-definite. This implies that for a positive cycle Y with $(Y^2) = 0$ we certainly have $\operatorname{supp}(Y) = X$. For, if not, there is an i with X_i not a component of Y but $(Y \cdot X_i) > 0$. Hence $(NY + X_i)^2 = 2N(Y \cdot X_i) + (X_i^2)$ is positive for large N, which is a contradiction. Choose the integral multiples aY, bZ as above. Then $(aY - bZ)^2 = 0$, and since the support is smaller that X, we must have $aY - bZ = 0$, i.e., $Y = b/aZ$, but this contradicts the assumption that (Z^2) is negative.

We revert to the original notation. Consider positive cycles $Z = \sum r_i X_i$ which have the property considered in the proposition, viz.

$$(*) \qquad (Z \cdot X_i) \leqq 0 \text{ for all } i.$$

If $Z^1 = \sum r_i^1 X_i$ and $Z^2 = \sum r_i^2 X_i$ have his property, so does $Z = \inf(Z^1, Z^2) = \sum r_i X_i$ with $r_i = \inf(r_i^1, r_i^2)$, and moreover Z is positive. In fact if say $r_j^1 \leqq r_j^2$, we have

$$(Z \cdot X_j) = r_j^1(X_j^2) + \sum_{i \neq j} r_i(X_i \cdot X_j) \leqq r_j^1(X_j^2) + \sum_{i \neq j} r_i^1(X_1 \cdot X_j)$$
$$= (Z^1 \cdot X_j) \leqq 0.$$

One sees easily that since Z^ν is positive and X connected, $Z^\nu \geqq X = \sum 1 \cdot X_i$. Hence also $Z \geqq X > 0$.

It follows that there is a unique smallest positive cycle Z satisfying (*). We call this Z the *fundamental cycle* of X.

THEOREM 3. *Let Z be the fundamental cycle of X. Then $p(Z) \geqq 0$ and $p(Z) = 0$ if and only if the singularity of $\bar{V}$ is rational.*

Proof. Let Y be a positive cycle with $Y \not\geqq Z$. Then $(Y \cdot X_i) > 0$ for some i, and one has by [5],

$$p(Y + X_i) = p(Y) + p(X_i) + (Y \cdot X_i) - 1.$$

Since $p(X_i) \geqq 0$,

$$(**) \qquad p(Y + X_i) \geqq p(Y).$$

If $0 < Y < Z$, then the multiplicity of this X_i in Y is necessarily less than the multiplicity of X_i in Z, i.e., $Y + X_i \leqq Z$. Since $p(X_1) \geqq 0$ to begin with, (**) shows that $p(Z) \geqq 0$ by induction.

Now suppose $p(Z) = 0$, and let Y_1 be any positive cycle. Define Y_ν inductively as follows:

(a) If $Y_{\nu-1} \geqq Z$, set $Y_\nu = Y_{\nu-1} - Z \geqq 0$.

(b) If $Y_{\nu-1} \not\geqq Z$, then by definition of Z, $(Y_{\nu-1} \cdot X_i) > 0$ for some i. Choose such an X_i with smallest possible multiplicity in $Y_{\nu-1}$ and set $Y_\nu = Y_{\nu-1} + X_i$.

Stop the process when $Y = 0$. Calculating the genus, we find in case (a),

$$\begin{aligned} p(Y_{\nu-1}) = p(Y_\nu + Z) &= p(Y_\nu) + p(Z) + (Z \cdot Y_\nu) - 1 \\ &\leqq p(Y_\nu) - 1 \text{ by } (*). \end{aligned}$$

Hence in case (a), $p(Y_\nu) > p(Y_{\nu-1})$. In case (b), $p(Y_\nu) \geqq p(Y_{\nu-1})$ by (**). Now one sees easily that a step of type (b) can not be repeated indefinitely without reaching a stage where $Y \geqq Z$. Moreover the function $p(Y)$ is bounded above since it is a quadratic function of the coefficients s_i of $Y = \sum s_i X_i$ whose quadratic term is $\frac{1}{2} \sum s_i s_j (X_i \cdot X_j)$, and since $\| (X_i \cdot X_j) \|$ is negative definite. Therefore, since each application of step (a) increases the value of the function p the process must result in $Y_\nu = 0$ for some ν, and then $Y_\nu = Y_{\nu-1} - Z$. Hence

$$p(Y_1) \leqq p(Y_{\nu-1}) < p(Y_\nu) = p(0) = 1,$$

i.e., $p(Y_1) \leqq 0$, and the theorem follows from Proposition (1).

Let $\mathfrak{m}$ be the maximal ideal of the affine ring A of $\bar{V}$:

THEOREM 4. *Suppose $\bar{V}$ has a rational singularity, and let Z be the fundamental cycle of X. Then for all integers n,*

$$V \times_{\bar{V}} \operatorname{Spec}(A/\mathfrak{m}^n) = nZ.$$

Moreover,

$$H^0(nZ, \mathcal{O}_{nZ}) = A/\mathfrak{m}^n,$$

and

$$\dim_K \mathfrak{m}^n/\mathfrak{m}^{n+1} = -n(Z^2) + 1.$$

Proof. We may replace A by its henselisation for this theorem, i.e., we may suppose that A is henselian. Since the sheaf of ideals $\mathfrak{J}(nZ) = (\mathfrak{J}(Z))^n$, the whole point for the first assertion is to show that

$$V \times_{\bar{V}} \operatorname{Spec} k = Z,$$

i.e., that $\mathfrak{m} \cdot \mathcal{O}_V = \mathfrak{J}(Z)$. Let $0 \neq f \in \mathfrak{m}$ and consider the divisor (f) of f on V. We can write it in the form

$$(f) = D + Y$$

where $0 < Y = \sum s_i X_i$ has support in X and $0 < D$ does not involve any X_i. We have

$$((f) \cdot X_i) = 0 \text{ and } (D \cdot X_i) \geqq 0 \text{ for all } i,$$

hence

$$(Y \cdot X_i) \leqq 0 \text{ for all } i,$$

and so

$$Y \geqq Z.$$

Therefore $\mathfrak{m} \cdot \mathcal{O}_V \subseteq \mathfrak{J}(Z)$.

It remains to show that for each point $P \in Z$, there exists a function $f \in \mathfrak{m}$ whose divisor is equal to Z locally at P. Choose a positive divisor D not passing through P and such that

$$(D \cdot X_i) = -(Z \cdot X_i) \text{ for all } i.$$

Since A is henselian, such divisors can be constructed at will. It suffices to find a rational function f on V such that $(f) = D + Z$. For, since A is normal and $(f) > 0$, it will follow that $f \in \mathfrak{m}$. This is equivalent with finding a nowhere zero section of the invertible sheaf $\mathcal{L} = \mathcal{L}(-(D+Z))$, i.e., to showing $\mathcal{L} \approx \mathcal{O}_V$. We will show that there is a "formal" section, i.e., an element of $\varprojlim_n H^0(nZ, \mathcal{L} \otimes_{\mathcal{O}_V} \mathcal{O}_{nZ})$, which induces a section s_0 of $H^0(Z, \mathcal{L} \otimes_{\mathcal{O}_V} \mathcal{O}_Z)$ which is nowhere zero. This is sufficient since by the theorem on holomorphic functions ([4], (4.2.1)) such a section can be approximated by a section s of $H^0(V, \mathcal{L})$ sufficiently close so that the section induced by s in $H^0(Z, \mathcal{L} \otimes_{\mathcal{O}_V} \mathcal{O}_Z)$ is s_0. Then necessarily s will be nowhere zero.

Now since $\bar{V}$ has a rational singularity, invertible sheaves on any nZ are classified by their degree ([1], (1.7)). Since $((D+Z)\cdot X_i)=0$ for all i the degree of $\mathcal{L}\otimes\mathcal{O}_{nZ}$ is zero and so $\mathcal{L}\otimes\mathcal{O}_{nZ}\approx\mathcal{O}_{nZ}$ for all n. Therefore it suffices to lift sections of $\mathcal{O}_Z$ successively to sections of $\mathcal{O}_{nZ}$ for each n, i.e., to show that

$$(*)\qquad H^0(Z,\mathcal{O}_{(n+1)Z})\rightarrow H^0(Z,\mathcal{O}_{nZ})\text{ is surjective}$$

for each n. Let $\mathcal{J}_n=\ker(\mathcal{O}_{(n+1)Z}\rightarrow\mathcal{O}_{nZ})$. We want to show that $H^1(Z,\mathcal{J}_n)=0$. Now one sees easily that

$$\mathcal{J}_n\approx\mathcal{I}(nZ)\otimes_{\mathcal{O}_V}\mathcal{O}_Z,$$

so $\mathcal{J}_n$ is an invertible sheaf. Its degree $d=(d_1,\cdots,d_n)$ is given by

$$d_i=(-nZ\cdot X_i)\geqq 0.$$

Hence the result is implied by the following simple lemma, the proof of which is in [5].

Lemma 5. *If* $\mathcal{F}$ *is an invertible sheaf on a cycle* $Z=\sum r_iX_i$ *satisfying the conditions of* ([1], (1.7)) *whose degree* $d=(d_1,\cdots,d_n)$ *has* $d_i\geqq 0$, *for all* i, *then*

$$\text{length}\,(H^0(Z,\mathcal{F}))=\sum r_id_i+1$$

and

$$H^1(Z,\mathcal{F})=0.$$

Moreover, because of $(*)$, it follows that the map $H^0(nZ,\mathcal{O}_{nZ})\rightarrow A/\mathfrak{m}^n$ is surjective. To show that it is injective, we have to show that if $g\in A$ has the property that $(g)=nZ+D$ with $D\geqq 0$, then $g\in\mathfrak{m}^n$. Suppose first that D meets each X_i transversally. Since A is henselian, the intersections of D with X are in one to one correspondence with the connected components of the support of D. Therefore we may decompose D (non-uniquely) into a sum $D=D_1+\cdots+D_n$ where $(D_\nu\cdot X_i)=1/n(D\cdot X_i)=-(Z\cdot X_i)$ for each i and ν. As above, there are rational functions f_ν such that $(f_\nu)=Z+D_\nu$. Then $g=f_1\cdots f_nu$ for some unit u of A, hence $g\in\mathfrak{m}^n$. Now if g is arbitrary and if $g_1\in\mathfrak{m}^n$ is chosen so as to have the transversality property above, then it is easily seen that $g_2=g+ag_1$ also has the transversality property for suitable $a\in A$, hence $g_2\in\mathfrak{m}^n$, and so $g\in\mathfrak{m}^n$.

It follows that $H^0(Z,\mathcal{J}_n)=\mathfrak{m}^n/\mathfrak{m}^{n+1}$. By the above lemma,

$$\begin{aligned}\dim_k\mathfrak{m}^n/\mathfrak{m}^{n+1}&=\text{length}\,H^0(Z,\mathcal{J}_n)\\&=\sum r_id_i+1=-n\sum r_i(X_i\cdot nZ)+1\\&=-n(Z\cdot Z)+1.\end{aligned}$$

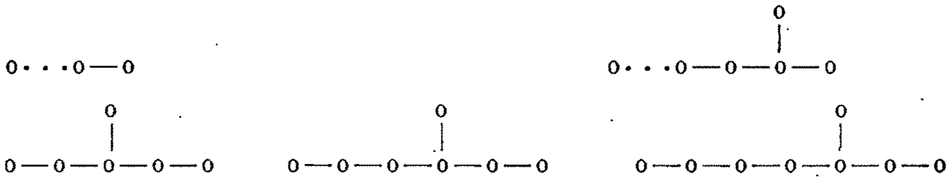
This completes the proof.

COROLLARY 6. *If $\bar{V}$ has a rational singularity then the multiplicity of A is $-(Z^2)$ where Z is the fundamental cycle, and the dimension of the Zariski tangent space of A is $-(Z^2)+1$.*

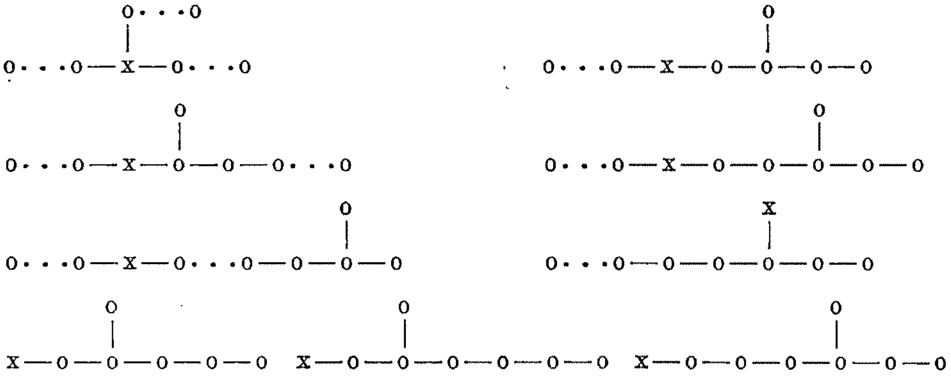
This follows immediately from the definition of multiplicity. In fact, the Hilbert polynomial $P(n)$ for the pair $(A,\mathfrak{m})$ is

$$P(x)=\tfrac{1}{2}m(x^2-x)+x.$$

The fact stated in the introduction that the configurations listed by Du Val (figure) arise only from double points can now be verified easily.



Rational double points. ($o^2=-2$)



Rational triple points. ($o^2=-2$, $x^2=-3$)

FIGURE

Combining (3) and (6) with the criterion of contractability of ([1], (2.7)) one obtains the generalization to reducible curves of Castelnuovo's criterion [2]:

COROLLARY 7. *Let $X=\cup X_i$, $i=1,\cdots,n$, be a connected complete curve on a nonsingular surface S over k. Then X is an exceptional curve of the first kind if and only if $\|(X_i\cdot X_j)\|$ is negative definite and the fundamental cycle $Z=\sum r_iX_i$ of X satisfies*

$$p(Z)=0 \text{ and } (Z^2)=-1$$

The figure below depicts the configurations which arise as "minimal" resolutions (i. e., without exceptional curves of the first kind) from rational double or triple points. They are the only ones theoretically possible according to (2) and (3), and they do all occur. In the figure a letter o denotes a nonsingular rational curve with self-intersection -2, and a letter x denotes a nonsingular rational curve with self-intersection -3. The lines — represent transversal intersections—the only ones possible, and the notation o· · ·o— stands for a chain of curves of arbitrary length, possibly length zero.

As an aid for the verification of the figure, note that in a triple point the fundamental cycle must have $p(Z)=0$ and $(Z^2)=-3$, hence $(Z\cdot K)=+1$. Therefore each configuration must contain exactly one curve x and it must have multiplicity 1 in Z.

REFERENCES.

[1] M. Artin, "Some numerical criteria for contractability of curves on algebraic surfaces," *American Journal of Mathematics*, vol. 84 (1962), pp. 485-496.

[2] G. Castelnuovo and F. Enriques, "Sopra alcune questioni fondamentali nella teoria delle superficie algebriche," *Annali di Matematica pura ed applicata* III s. vol. 6 (1891), pp. 165-225.

[3] P. Du Val, "On isolated singularities of surfaces which do not affect the conditions of adjunction" (Part I), *Proceedings of the Cambridge Philosophical Society*, vol. 30 (1933-34), pp. 483-491.

[4] A. Grothendieck, *Éléments de géométrie algébrique*, Ch. III, Institut des Hautes Études Scientifiques, Publications Mathématiques No. 11 (1961).

[5] S. Lichtenbaum, *Curves over discrete valuation rings*, Doctoral dissertation, Harvard (1964) (forthcoming).

[6] D. Mumford, *The topology of normal singularities of an algebraic surface and a criterion for simplicity*, Institut des Hautes Études Scientifiques, Publications Mathématiques No. 9 (1961).

[7] ———, Appendix to a paper of O. Zariski "The theorem of Riemann-Roch for high multiples of an effective divisor on an algebraic surface," *Annals of Mathematics*, 2nd Ser., vol. 76 (1962), pp. 560-615.

[8] O. Zariski, *Algebraic Surfaces*, Ergebnisse der Mathematik, Bd. 5 Berlin (1935).

[9] ———, *Introduction to the problem of minimal models in the theory of algebraic surfaces*, Publications of the Mathematical Society of Japan, No. 4, Tokyo (1958).

[10] ———, "The theorem of Riemann-Roch for high multiples of an effective divisor on an algebraic surface," *Annals of Mathematics*, 2nd Ser., vol. 76 (1962), pp. 560-615.

ON THE CLASSIFICATION OF DIFFERENTIABLE $SO(n)$ ACTIONS ON SIMPLY CONNECTED π-MANIFOLDS.

By Wu-yi Hsiang.

In the theory of topological transformation groups the classification problem is central. It is, of course, quite difficult. For example, mathematicians have studied the special case of Z_2 actions on the n-sphere for more than twenty years, yet only the cases $n=2,3$ have been successfully solved [10, 18]. We confine ourselves to the cases of compact Lie groups acting differentiably on smooth manifolds. This enables us to make effective use of both differential topology and Lie group theory. On the other hand, there are examples of Bing and Montgomery [2, 15] to show that both the differentiability and the compactness of the group are essential.

In this paper, we succeed in classifying all differentiable actions of $SO(n)$ on some interesting manifolds such as spheres, euclidean spaces and π-manifolds with dimension lower than $2n-1$. We note that $\dim(SO(n)) = \frac{1}{2}n(n-1)$ is comparatively high with respect to the dimension of the manifold, $m<2n-1$; this corresponds to some kind of regularity that the "symmetry" of the "structure" is quite rich. The author was somehow suggested by the following classical theorem in Riemannian geometry: "The highest dimension of the isometry group of a Riemannian n-manifold M^n is $\frac{1}{2}n(n+1)$, and the maximal possibility occurs only when $M^n=S^n$ or P^n with the usual Riemannian structure; in this case, the isometry group is $O(n+1)$."

This motivates us to study large group actions rather than to study small group actions such as finite groups actions and $SO(3)$ actions. One of the main technical reasons to restrict the dimension $m<2n-1$ is the following Lie group theoretical fact:

"Let $H\subseteq SO(n)$ $(n\geqq 9)$ be a closed subgroup with codimensions less than $2n-1$, then $H_0=SO(n-1)$, $SO(n-2)$ or $SO(n-2)\times SO(2)$." Hence, in case $SO(n)$ acts on M^m with $m<2n-1$, the orbit structure will be considerably simpler. In fact, with only four exceptional cases, it is proved by combining slices structure and vector bundle theory that there are only orbits of the type of S^{n-1} and fixed points for π-manifolds.

Based on the above local results on orbit structures we show (by blowing

---

Received October 25, 1964.

up along the set of fixed points) that the associated orbit spaces are naturally smooth manifolds with boundary and that two actions are differentiably equivalent if and only if their associated orbit spaces are diffeomorphic. In case that the manifold is an m-sphere, it is proved that the orbit space is a compact contractible $(m-n+1)$-manifold, and any compact contractible $(m-n+1)$-manifold is the orbit space of a suitable $SO(n)$ action on S^m. Similar topological characterization for the cases $M=D^m$ and $M=R^m$ are also given.

A number of geometric consequences and applications are deduced. For example, we show that there is no differentiable action of $SO(n)$ on an exotic sphere of dimension less than $2n-1$. This shows that the differentiable symmetry of exotic spheres are far poorer than those of ordinary spheres.

Acknowledgment

The author thanks his advisors Professors J. C. Moore and N. E. Steenrod for the invaluable guidance they gave him. Thanks also go to Professor J. Milnor and Professor D. Montgomery for their encouragements and several useful comments.

Classification theorems. To be precise, let us first define the category to be used.

Let $\mathcal{G}$ be the category of differentiable G-spaces and equivariant differentiable maps, namely, an object in $\mathcal{G}$ is a triple (M,G,Φ) where M is a smooth manifold; G is a compact Lie group and $\Phi\colon G\times M\to M$ is a smooth action map in the usual sense; a map in $\mathcal{G}$ is a pair

$$(M_1,G_1,\Phi_1)\xrightarrow{(f,h)}(M_2,G_2,\Phi_2)$$

where $f\colon M_1\to M_2$ is a smooth map and $h\colon G_1\to G_2$ is a homomorphism of Lie groups and the following diagram is commutative;

$$\begin{tikzcd}
G_1\times M_1 \arrow[r, "\Phi_1"] \arrow[d, "h\times f"] & M_1 \arrow[d, "f"] \\
G_2\times M_2 \arrow[r, "\Phi_2"] & M_2.
\end{tikzcd}$$

To any given triple (M,G,Φ), there is associated naturally a Hausdorff space X called the orbit space of (M,G,Φ); and to any map (f,h) in $\mathcal{G}$, there is induced a natural map $\bar{f}$ between the two orbit spaces. It is clear that assigning to objects and maps in $\mathcal{G}$ their associated orbit spaces and maps, respectively, is a "functor" from $\mathcal{G}$ to the category of spaces and continuous maps.

For a given manifold M and a given group G, we denote the set of equivalence classes of *non-trivial* triples $\{(M, G, \Phi_i)\}$ in $\mathcal{B}$ by $\Gamma(M, G)$.

1. Constructions. The following are some constructions that will make the contents of the classification theorems clearer and the proofs of them appear more natural.

Construction 1. Let $n \geqq 4$ and X^k be a simply connected π-manifold (possibly with boundary). Let $(D^n, SO(n), \Phi_0)$ be the unit n-disc with $SO(n)$ acting orthogonally.

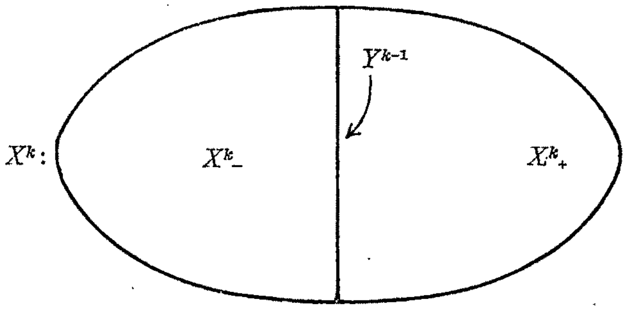
We then have an $SO(n)$-space $(D^n \times X^k, SO(n), \Phi_0 \times \mathrm{id})$ (with corners equivariantly smoothed). If we restrict the action to the smoothed boundary, we have an $SO(n)$-space $(M^{n+k-1}, SO(n), \Phi_1)$ with $M^{n+k-1} = \partial(D^n \times X^k)$ (with corners smoothed) and $\Phi_1 = \Phi_0 \times \mathrm{id} \mid \partial(D^n \times X^k)$.

It is easy to check that M^{n+k-1} is again a simply connected π-manifold and the orbit space of $(M^{n+k-1}, SO(n), \Phi_1)$ is diffeomorphic to X^k while the image of the set of fixed points, F, corresponds exactly to $\partial(X^k)$.

Construction 2. If in particular, X^k is a compact contractible manifold, then we will show that the $SO(n)$-space M^{n+k-1} constructed above is actually an $(n+k-1)$-sphere.

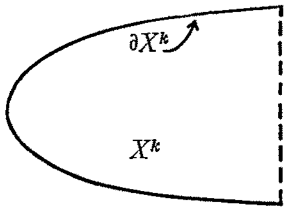
Proof. Since X^k is now contractible, we see easily that $D^n \times X^k$ is also compact contractible with simply connected boundary. If $k \leqq 2$, we have $X^k = D^k$ and we have trivially $D^n \times X^k = D^{n+k}$. If $k \geqq 3$, we have $n + k \geqq 7$, then by h-cobordism theorem, we again have $D^n \times X^k = D^{n+k}$ and hence $M^{n+k-1} = \partial(D^n \times X^k) = S^{n+k-1}$.

Construction 3. Let (X^k_+, Y^{k-1}) be a pair of compact contractible manifolds with $Y^{k-1} \subseteq \partial X^k_+$, and with $\frac{\pi}{2}$-corner along ∂Y^{k-1}. Then we may double X^k_+ along Y^{k-1} to get a smooth compact contractible manifold $X^k = X^k_+ \bigcup_{Y^{k-1}} X^k_-$. By construction 2, we have an $SO(n)$-space $(S^{n+k-1}, SO(n), \Phi_1)$ with its space of orbits diffeomorphic to X^k.



Let $M^{n+k-1} = \pi^{-1}(X^k_+)$, then it is clear that $\partial M^{n+k-1}_+ = \pi^{-1}(Y^{k-1})$ and hence $\partial M^{n+k-1}_+ = S^{n+k-2} \subseteq S^{n+k-1}$. Hence, we have that $M^{n+k-1}_+ = D^{n+k-1}$, that is, we have thus constructed an $SO(n)$-action on (D^{n+k-1}, S^{n+k-2}) such that their orbit space is the given pair (X^k_+, Y^{k-1}).

Construction 4. Let X^k be a non-compact contractible manifold with boundary ∂X^k such that X^k is simply connected at infinity and ∂X^k is acyclic.



Then following construction 1, we have an $SO(n)$-space M^{n+k-1} which is open, contractible and simply connected at infinity. By Stallings [19], $M^{n+k-1} = R^{n+k-1}$. Thus, to any given such X^k, there exists an $SO(n)$-action on R^{n+k-1} with its orbit space diffeomorphic to X^k.

2. Blowing up argument. The following is an intermediate stage for the classification theorems that we are going to prove later. It involves to replace a G-space with only two orbit types by another G-space of only one orbit type, particularly for the case that there are only orbits of sphere type and fixed points.

THEOREM I. *Let M^m be a G-space with orbits of S^{n-1}-type and fixed points only, and F be the set of fixed points.*

Then, by blowing up along F, we have a G-space with boundary $(\widetilde{M}^m, \partial\widetilde{M}^m)$ and an equivariant relative diffeomorphism $(\widetilde{M}^m, \partial\widetilde{M}^m) \xrightarrow{h} (M^m, F)$ such that M^m is a G-space with only orbits of S^{n-1}-type and the following diagram is commutative,

$$\begin{tikzcd}
(\widetilde{M}^n, \partial\widetilde{M}^m) \arrow[rr, "h"] \arrow[dr] & & (M^m, F) \arrow[dl] \\
& (X^k, \partial X^k) &
\end{tikzcd}$$

where the common orbit space X^k is a manifold with boundary $\partial X^k = \pi(F)$, and $k = m - n + 1$.

Proof. Since D^n can be viewed as a smooth cone over S^{n-1}, and the blowing up process of a cone over S^{n-1} at the vertex is simply replacing the cone by a cylinder $S^{n-1} \times I$.

In our case, we first assume that G acts as a group of isometries with respect to a suitably chosen metric on M. Then the blowing up along F is just to replace an *invariant normal disc-bundle equivariantly* by the associated cylinder bundle $\nu \times I$, where ν is the normal sphere bundle of F in M^m. Hence the theorem follows.

COROLLARY 1. *If M^m is simply connected and $n \geqq 2$, then X^k is also simply connected.*[1]

Proof. Since M^m is a S^{n-1} fibration over X^k and S^{n-1} is connected, we have

$$\begin{tikzcd}
\pi_1(\widetilde{M}^m) \arrow[rr, "\text{epi}"] \arrow[dr, "\tilde{h}"] & & \pi_1(X^k) \arrow[r] & 0 \\
& \pi_1(M^m) \arrow[ur] & &
\end{tikzcd}$$

Hence, $\pi_1(M^m) \xrightarrow{\text{epi}} \pi_1(X^k)$ and $\pi_1(X^k) = 0$.

COROLLARY 2. *If $G = SO(n)$ and M^m is simply connected, then $(M^m, SO(n), \Phi)$ is differentiably equivalent to $(\partial(D^n \times X^k), SO(n), \Phi_1)$ given in construction* 1.

Proof. By Borel [5], the structure group of the fibration $S^{n-1} \to \widetilde{M}^m \to X^k$ can be reduced to $N(SO(n-1))/SO(n-1)$, which is Z_2. If X^k is simply connected, we see easily that $\widetilde{M}^m = S^{n-1} \times X^k$ and the corollary follows.

3. Statements of the classification theorems.

THEOREM II. *Let $(M^m, SO(n), \Phi)$ be an $SO(n)$-space such that $n \geqq 9$, $m < 2n - 3$ and M^m is a simply connected π-manifold. Then its space of orbits X is an $(m - n + 1)$-dimensional simply connected π-manifold (possible with boundary) and $(M^m, SO(n), \Phi)$ is differentiably equivalent to the triple $(\partial(D^n \times X), SO(n), \Phi_1)$ given in the construction* 1. (*i.e. every $SO(n)$-space $(M^m, SO(n), \Phi)$ with the above condition is differentiably equivalent to one and only one example given in construction* 1.)

Remark. Theorem II is still true for $m = 2n - 3$ and $m = 2n - 2$ but with the following four exceptions:

---

[1] More general result may be shown by path lifting in slices. For example, see [14].

(i) $(M_1, SO(n), \Phi_1)$ with
$M_1 = V_{n,2} = SO(n)/SO(n-2)$ and $\Phi_1 \colon SO(n) \times V_{n,2} \to V_{n,2}$ is the usual map induced from multiplication.

(ii) $(M_2, SO(n), \Phi_2)$ with $M_2 = V_{n,2} \times R$. R the real line and $\Phi_2 = \Phi_1 \times \mathrm{id}$.

(iii) $(M_3, SO(n), \Phi_3)$ where $M_3 = \tau(S^{n-1})$ is the tangent vector bundle of S^{n-1} and Φ_3 is the natural action map induced from the only orthogonal action of $SO(n)$ on S^{n-1}.

(iv) $(M_4, SO(n), \Phi_4)$ where M_4 is the double of the associate disc bundle of $\tau(S^{n-1})$ and Φ_4 is the natural inherited action map.

THEOREM III. *Let* $(\Sigma^m, SO(n), \Phi)$ *be an* $SO(n)$*-space such that* $n \geqq 9$, $m < 2n-1$ *and* Σ^m *is a homotopy sphere. Then the associate orbit space* X *is an* $(m-n+1)$*-dimensional compact contractible manifold.*

Hence, by Theorem II, $(\Sigma^m, SO(n), \Phi) \sim (\partial(D^n \times X), SO(n), \Phi_1)$ *given in construction* 2, *and every differentiable* $SO(n)$ *action on m-homotopy sphere is differentiably equivalent to one and only example constructed in construction* 2.

THEOREM IV. *Let* $(D^m, SO(n), \Phi)$ *be a differentiable action of* $SO(n)$ *on* D^m, $n \geqq 9$, $m \leqq 2n-1$; *then the space of orbits is a pair* $(X^{m-n+1}{}_{+}, Y^{m-n})$ *described in construction* 3, *and* $(D^m, SO(n), \Phi)$ *is differentiably equivalent to the unique example constructed there from the given pair* $(X^{m-n+1}{}_{+}, Y^{m-n})$.

THEOREM V. *Given an* $SO(n)$ *action on* R^m, $(R^m, SO(n), \Phi)$ *then the orbit space* X^k *is a space described in construction* 4 *with* $k = m-n+1$, *and* $(R^m, SO(n), \Phi)$ *is differentiably equivalent to the unique example given in construction* 4 *with its orbit space* X^k $(m \leqq 2n-1)$.

4. Proofs of the classification theorems. The main step is to show that there are only orbits of S^{n-1}-type and stationary points. Then the theorems follow from the blowing up argument.

Let M^m be a π-manifold with a (non-trivial) differentiable $SO(n)$ action, $n \geqq 9$ and $m < 2n-1$. Let $x \in M^m$, we denote the isotropy subgroup of x by G_x, i.e., $G_x = \{g \in SO(n) \mid g \cdot x = x\}$. Then the orbit passing through x, $G(x)$, is a smooth embedding of $SO(n)/G_x$. Hence, we trivially have $\operatorname{codim} G_x = \dim SO(n)/G_x \leqq m < 2n-1$.

LEMMA 1. *Let* H *be a closed proper subgroup of* $SO(n)$, $n \geqq 9$ *and* $\operatorname{codim} H = \dim(SO(n)/H) < 2n-1$. *Then the identity component of* H, H_0, *is either* $SO(n-1)$, or $SO(n-2)$, *or* $SO(n-2) \times SO(2)$.

Proof. Much more general result may be proved by studying the representation $H_0 \subseteq SO(n)$. We may show that the representation is always reducible if the codimension of H_0 in $SO(n)$ is comparatively small. We refer to Dynkin [9] for details.

We observe that $H_0 \subseteq H \subseteq N(H_0)$ ($N(H_0)$ is the normalizer of H_0 in $SO(n)$). Hence, for the moment, the possible candidate for G_x is $SO(n)$, $SO(n-1)$, $N(SO(n-1))$, $SO(n-2)$, $SO(n-2) \times Z_h$, $SO(n-2) \times SO(2)$ and $N(SO(n-2) \times SO(2))$.

We will eliminate most of the possibility by means of vector bundle theory. Let us first recall some basic facts in topological transformation groups in terms of vector bundles.

(i) Let G be a compact group acting differentiably on a smooth manifold M. Then by averaging an arbitrarily given Riemannian metric on M, we may have a new metric such that G acts as a group of isometries.

(ii) Let $x \in M$, G_x be the isotropy subgroup leaving x fixed, and $G(x) \approx G/G_x$ be the orbit passing through x. Then $G(x)$ is a smooth compact submanifold, and G_x acts on a small normal disc of $G(x)$ at x orthogonally; we call it the normal representation of G_x at x and is denoted by ϕ_x.

(iii) Let ν be the normal vector bundle of $G(x) = G/G_x$. Then ν is the weakly associated vector bundle of the natural principal G_x-bundle $G_x \to G \to G/G_x$ via the normal representation ϕ_x. Namely $E(\nu) = G \times_{G_x} R^k$ where $k = \operatorname{codim} G(x)$ and G_x acting on R^k via ϕ_x. We note that G acts naturally on $E(\nu)$ as bundle mappings and we may choose small enough ϵ such that the exponential map maps the ϵ-disc bundle of ν equivariantly and diffeomorphically onto an invariant tubular neighborhood, $N(G(x), \epsilon)$, of $G(x)$. This is the vector bundle interpretation of the classical slice structure [14].

(iv) Let $H \subseteq G$ be a closed subgroup, we denote the set of all conjugate subgroup of H in G by (H). We introduce the following partial ordering relation "$\prec$" by defining $(H_1) \prec (H_2)$ if there exist $H_1 \in (H_1)$ and $H_2 \in (H_2)$ such that $H_1 \subseteq H_2$. If G acts differentiably on M, then the set of all possible isotropy subgroups $\{G_x \mid x \in M\}$ splits naturally into conjugate classes. If M is connected, then it was proved by Montgomery [15] that there always exists an absolute minimal, (H), among the conjugate classes in $\{G_x \mid x \in M\}$, moreover, the set of points

$$O_{(H)} = \{x \in M \mid G_x \in (H)\}$$

is an everywhere dense connected open submanifold. (H) is called the principal orbit type.

(v) Combine (iv) and (iii), we see that ϕ_x is trivial if and only if the orbit $G(x)$ is a principal orbit. Hence, we have the following lemma trivially.

LEMMA 2. *A necessary condition for a homogeneous space G/H to be a principal orbit of certain G action on M is that G/H may imbedded in M with trivial normal bundle. In particular, if M is stably parallelizable then the principal orbit must also be stably parallelizable.*[2]

Proof of Theorem II. It follows from Lemma 1 and Lemma 2 that the only possible candidate for principal orbit type is either $(SO(n-1))$ or $(SO(n-2))$; since they are the only stably parallelizable homogeneous spaces of $SO(n)$ of dimension less than $2n-1$, $n \geqq 9$.

Case I. If the principal isotropy subgroup is $SO(n-2)$, we note that $\dim(SO(n)/SO(n-2)) = 2n-3$. Hence this is possible only when $m = 2n-2$ or $2n-3$. If $m = 2n-3$, then we trivially have $M^m = V_{n,2} = SO(n)/SO(n-2)$ and the action is transitive. Suppose $m = 2n-2$, then the orbit space is one-dimensional. By (iii), we see that it is either the real line or a half line $[0, \infty)$; or is an interval $[0, 1]$. It is also a consequence of (iii) that the ending points are the image of singular orbits of S^{n-1}-type. From there, it is not difficult to see that the four exceptional cases mentioned in the Remark are the only possible cases with $SO(n-2)$ as principal isotropy subgroup.

Case II. The principal isotropy subgroup is $SO(n-1)$. Now, we only have to show that it is impossible to have a point $x \in M^m$ with $G_x = N(SO(n-1))$, namely, it is impossible to have

$$P^{n-1} = SO(n)/N(SO(n-1))$$

as an orbit. We show by contradiction.

Suppose there is a point $x \in M^m$ such that $G(x) = P^{n-1}$. Let ν be the normal vector bundle of $G(x)$ in M^m, and ϕ_x is the normal representation at x;

---

[2] It was also stated in a paper of Borel-Hirzebruch.

$$\phi_x\colon N(SO(n-1)) \to 0(m-n+1).$$

We claim that $SO(n-1) \subseteq N(SO(n-1))$ is the kernel of the representation ϕ_x. [Since $SO(n-1)$ is a normal subgroup as well as the principal isotropy subgroup of ϕ_x, by Montgomery [15], $SO(n-1) \subseteq N(SO(n-1))$ acts trivially on the normal disc.] Now, by the slice structure (iii), ν is weakly associated to

$$N(SO(n-1) \to SO(n) \to P^{n-1}$$

via ϕ_x. But we have shown that

$$\begin{tikzcd}
SO(n-1) \arrow[r] & N(SO(n-1)) \arrow[r, "\phi_x"] \arrow[dr] & O(m-n+1) \\
& & Z_2 \arrow[u, "\tilde{\phi}_x"']
\end{tikzcd}$$

i.e., ν is actually associated to the following principal bundle

$$Z_2 \to S^{n-1} \to P^{n-1}$$

via $\tilde{\phi}_x$. Let the image of the generator of Z_2 under $\tilde{\phi}_x$ be a diagonal matrix

$$\begin{pmatrix} -I_a & 0 \\ 0 & I_b \end{pmatrix}.$$

Then $\nu = a\xi + b\theta$, $(a+b = m-n+1)$ where ξ is the usual projective line bundle and θ is the trivial line bundle.

Since $\tau(M^m)$, the tangent bundle of M^m, is stably trivial, we have

$$\tau(P^{n-1}) \oplus \nu = 0 \text{ in } \widetilde{KO}(P^{n-1})$$

i.e., $$\tau(P^{n-1}) \oplus \theta + \nu = n\xi + a\xi + b\theta = (n+a)\xi = 0$$

According to the computation of F. Adams [1], $(n+a) = k2^{[n-1]/2}$ which is impossible for $(n+a) \leqq 2n-1$, $n > 9$.

Hence there are no orbits of P^{n-1}-type, the only remaining possibilities are S^{n-1} and stationary points. Now, Theorem I applies and Theorem II follows immediately.

Proof of Theorem III. It follows from Theorem II that the orbit space is a simply connected π-manifold (possibly with boundary) and two action with diffeomorphic orbit spaces are differentiably equivalent. Moreover, in construction 2, we show that every compact contractible manifold of $(m-n+1)$ dimension can be realized as the orbit space of suitable action. Hence, we only have to show that the orbit space $\Sigma^m/SO(n)$ is always contractible.

Since we know $X=\Sigma^m/SO(n)$ is simply connected, we need only to show that X is acyclic. Since Σ^m cannot be $S^{n-1}\times X$, we see that the set of stationary points $F\approx\partial X\neq\emptyset$. Hence $\dim F=m-n$, and by Bredon [6], we know that X is a cohomologically disc, and hence X is acyclic. q. e. d.

Proof of Theorem IV. Let $(D^m, SO(n), \Phi)$ be a nontrivial action of $SO(n)$ on D^m, $m\leqq 2n-1$, $n\geqq 9$. Let (X_+, Y) be the orbit space and the subspace of orbits in $\partial D^m=S^{m-1}$. We may double D^m equivariantly to get an action of $SO(n)$ on S^m. It is obvious that the orbit space of $(S^m, SO(n), \Phi'), X$, is the double of X_+ along a partial boundary Y.

It follows from Theorem III that both X and Y are contractible. By Mayer-Victoris sequence, we see easily that X_+ is acyclic and hence contractible. The other parts have already proved in Theorem II and construction 3.

Proof of Theorem V. Again, we only have to show that the orbit space is contractible, simply connected at infinity and having acyclic boundary ∂X.

First, we see that the set of stationary points F is non-empty and of highest dimension. Then by one point compactification, we get a topological action of $SO(n)$ on S^m with one added fixed point. By Bredon [6], we see that the one point compactification of X is a cohomological disc, and everything follows almost immediately.

5. Geometric consequences and applications of the classification theorems. The following are some of the consequences and applications of those classification theorems. Of course, there may be more, we only try to deduce those directly related ones.

THEOREM VI. *Let Σ^m be an exotic sphere of dimension m, $m\geqq 17$; then there is no non-trivial differentiable action of $SO(n)$, $n\geqq(m+2)/2$, on Σ^m.*

Proof. Suppose the contrary, we have by Theorem III, the orbit space is a compact contractible manifold X. On the other hand, we may construct a differentiable action on differentiable sphere S^m with X as orbit space. By Theorem III, these two are differentiable equivalent i. e. the given exotic sphere is not exotic, which is a contradiction. q. e. d.

The geometric significance of the above result is that the differentiable homogenuity of exotic spheres are far poorer than that of the ordinary differentiable spheres. It is interesting, maybe just for curiosity, to see what's the biggest possible group that may act differentiably on an exotic sphere. The author and his brother have constructed infinite S^1 and S^3 differentiable action on certain exotic 11-spheres. But we believe this is far from being big enough. On the other hand, combine the argument in this paper and a recent paper of Bredon [7, 8], when $m = 4k - 1$, we may replace the inequality $n \geqq (m+2)/2$ by $n \geqq (m+1)/2$. While in the case $m = 4k+1$, Bredon give infinite may action on homotopy spheres but he cannot determine whether there are exotic spheres among them or not. If there are exotic ones, then in some sense, our theorem is the best possible negative result.[3]

THEOREM VII. *Let W^m be a real (or complex, or quaternion) Stiefel manifold* of dimension $m < 2n - 2$, $n \geqq 9$ and $W^m \neq S^m$ or $SO(n)/SO(n-2)$. Then there is no non-trivial $SO(n)$ action on W^m.*

Proof. By Theorem II, if there exists a non-trivial $SO(n)$-action on W^m, let X be the associate orbit space. Then, we have that X is a manifold with boundary of dimension $k = m - n + 1 \leqq n - 2$, and $W^m = \partial(D^n \times X) = (S^{n-1} \times X) \cup (D^n \times \partial X)$.

Now $H^p(W^m, D^n \times \partial X) = H^p(S^{n-1} \times X, S^{n-1} \times \partial X)$. We consider the projection $\pi\colon D^n \times X \to X$ restrict to $W^m = \partial(D^n \times X)$ we have

$$\pi\colon (W^m, D^n \times \partial X) \to (X, \partial X).$$

Hence, we have

$$\begin{tikzcd}
{} \arrow[r] & H^p(\partial X) \arrow[r, "\partial"] \arrow[d, "\approx"] & H^{p+1}(X,\partial X) \arrow[r] \arrow[d, "\pi^*"] & H^{p+1}(X) \arrow[r] \arrow[d, "\pi^*"] & H^{p+1}(\partial X) \arrow[r, "\partial"] \arrow[d, "\pi^*", "\approx"'] & {} \\
{} \arrow[r] & H^p(D^n \times \partial X) \arrow[r, "\partial"] & H^{p+1}(W^m, D^n \times \partial X) \arrow[r] & H^{p+1}(W^m) \arrow[r] & H^{p+1}(D^n \times \partial X) \arrow[r, "\partial"] & {} \\
 & & H^{p+1}(S^{n-1} \times X, S^{n-1} \times \partial X) \arrow[u, "\approx"'] & & &
\end{tikzcd}$$

for $i < n-1$, $H^i(X, \partial X) \xrightarrow{\pi^*} H^i(W^m, D^n \times \partial X) \xrightarrow{\approx} H^i(S^{n-1} \times X, S^{n-1} \times \partial X)$

---

[3] According to Bredon, that he is now able to show that his example M_1^9 is an exotic sphere. Hence, our negative results is best possible for $m = 9$.

* Since Stiefel manifolds are principal orbit types of suitable representation of $SO(n)$ or $U(n)$, $Sp(n)$ respectively, we see easily that they are π-manifolds.

is an isomorphism, hence by five lemma, we see that $H^p(X) \xrightarrow{\pi^*} H^p(W^m)$ is an isomorphism for $p \leqq n-2$. This fact implies that the fundamental class (with Z or Z_2 as coefficients) of W^m cai *never be expressed* as the cup product of elements of dimensions less than $(n-1)$, which is clearly a contradiction to the cohomology ring structure of Steifel manifolds different from $SO(n)/SO(n-2)$ and S^m. q.e.d.

This is the first attempt to ask the question: What is the biggest (in dimension) compact Lie group that may act differentiably on a given space, especially those nice spaces such as homogeneous spaces. This question is closely related to the isometry group of a Riemannian manifold. For example, to a given manifold, we may equip it with infinite many Riemannian metrics, which one is the most natural? Suppose there is a reasonable answer, we would like to see that the natural one has the richest "symmetry," i.e., the full isometry group is the biggest in certain sense! For spheres and projective spaces, the classical theorem, that mentioned in the introduction gives us a perfect answer, but for other nice spaces such as homogeneous spaces, this question is even not well formulated yet! It seems to the author that this is one of the important and sensible questions that we may look after.

We recall that there exist compact contractible k-manifolds X^k for $k \geqq 4$ with $\pi_1(\partial X^k) \neq 0$. Let "#" denote the "connected sum" operation (along the boundary if the boundary of the two manifolds are both non-empty). If we let $X_i^k = X^k \# \cdots \# X^k$ (i-th sum), we easily get infinite many contractible manifold X_i^k with $\pi_1(X_i^k) \neq \pi_1(X_j^k)$ for $i \neq j$.

Hence, as a direct consequence of Theorem III, IV, we have infinite many differentiable $SO(n)$-action on S^m or D^m if $m-n \geqq 3$. Actually, the following more general assertion is true.

THEOREM VIII. *Let M^m be a smooth manifold and G be a compact Lie group. If there exists a differentiable G action on M with* $\dim F \geqq 3$, *F is the set of stationary points. Then there exists infinite many different G actions on M^m.*

Proof. Let $x \in F \not\subseteq M$ be a stationary point such that the dimension, k, of the component of x is larger than or equal to 3. By Bochner theorem, the action is orthogonal around x. This gives us a representation ϕ_x of G into $O(m)$. Since the fixed point set is k-dimensional around x, we see easily that there are exactly k trivial copies in ϕ_x. Let $\tilde{\phi}_x = \phi_x - k$, then $\tilde{\phi}_x$ is the normal representation without trivial copies, $\dim \tilde{\phi}_x = m-k$. By construction 2 and the above remark, there exist infinite many $SO(m-k)$

actions on S^m with their fixed point sets, say F_i, all different (in fact $\pi_1(F_i) \neq \pi_1(F_j)$ for $i \neq j$). We let G act on S^m via the representation $\tilde{\phi}_x$. Then we have countable many G-spaces (S^m, G, Φ_i) with their fixed point sets being F_i.

We observe that the normal representation without trivial copies, $\tilde{\phi}_{y_i}$, are always the same as $\tilde{\phi}_x$ for any given y_i in F_i, $(i = 1, 2, \cdots, n, \cdots)$. Hence, we may construct the "equivariant connected sum" of (M^m, G, Φ) with (S^m, G, Φ_i) along x and y_i. This is possible because the two G-space have the same linear action along x and y_i. This is possible because the two G-space have the same linear action along x and y_i.

It is clear that the fixed point sets of $(M^m, G, \Phi) \# (S^m, G, \Phi_i)$ are $F \# F_i$. Now $\pi_1(F_i) \neq \pi_1(F_j)$ implies $\pi_1(F \# F_i) \neq \pi_1(F \# F_j)$. Hence

$$(M^m, G, \Phi) \# (S^m, G, \Phi_i) = (M^m, G, \Phi \# \Phi_i)$$

are all different (topologically non-equivalent) and the theorem is proved. q. e. d.

COROLLARY 3. *Let* $n \geqq 9$ *and* $2n - 3 > m \geqq n + 2$ *and* M^n *be a simply connected* π*-manifold such that* M^m *is not a product of* $(n-1)$*-sphere with some other manifold. Then either there is no non-trivial* $SO(n)$*-action on* M^m *or there are infinitely many non-equivalent* $SO(n)$*-action on* M^m.

Proof. Suppose there is a non-trivial $SO(n)$-action on M^m, let X be the orbit space. Then by Theorem II $M^m = \partial(D^n \times X)$ and the fixed point set $F = O \times \partial X$. Since $M \neq S^{n-1} \times X$, we see that $F \approx \partial X$ is non-empty and of dimension $\geqq 3$. Hence the corollary follows from Theorem VIII.

COROLLARY 4. *Let* $m - n \geqq 3$ *then there exist infinite many topologically non-equivalent differentiable* $SO(n)$ *actions on* P^m*, the m-th projective space.*

Proof. Since we may easily exhibit an orthogonal $SO(n)$ action on P^m with its fixed point set diffeomorphic to P^{m-n}, and Theorem VIII applies.

THEOREM IX. *For* $n + 4 \leqq m \leqq 2n - 2$*, then an* $SO(n)$*-action on* S^m *is orthogonal if and only if the fixed point set* F *is simply connected.*

Proof. By Theorem III, the orbit space X^{m-n+1} is a compact contractible manifold with $\partial X \approx F$. If F is simply connected we have by h-cobordism theorem, X is a disc and hence, again by Theorem III, the action is orthogonal.

THEOREM X. *In case* $n + 5 \leqq m \leqq 2n - 1$*, then an* $SO(n)$*-action on* D^m *is orthogonal if and only if the fixed point set* F *is a disc.*

Proof. Similarly by Theorem IV and h-cobordism.

In the remaining of this paper, we consider some imbedding and extension problem in $\mathscr{C}$-category.

We observe that under equivalent connected sum operation at fixed points, we may "add two actions on S^m or D^m" to get a new action. Moreover, it is a consequence (highly non-trivial) of the classification Theorem III, IV that this operation is "well defined." Hence the sets $\Gamma(S^m, SO(n))$ and $\Gamma(D^{m+1}, SO(n))$, $m < 2n-1$, are commutative and associative semi-groups under connected sum operation. Actually, by Theorem II, IV, the "orbit space functor" maps the above semi-groups to the semi-group of compact contractible $(m-n+1)$-manifolds and the semi-group of pairs of compact contractible manifolds (under suitable connected sum operation).

Given a triple $\{(M, G, \Phi)\} \in \Gamma(M, G)$ there naturally associate a triple $(\partial M, G, \Phi \mid G \times \partial M) \in \Gamma(\partial M, G)$ and this gives a boundary operation

$$\partial\colon \Gamma(M, G) \to \Gamma(\partial M, G).$$

Let $h\colon G_1 \to G_2$ be a homomorphism of compact Lie groups and $\{(M, G_2, \Phi)\} \in \Gamma(M, G_2)$ then there natural induces a triple

$$\{(M, G_1, \Phi \circ (h \times \mathrm{id}))\} \in \Gamma(M, G_1)$$

and defines a map (functorial and contravariant)

$$h^*\colon \Gamma(M, G_2) \to \Gamma(M, G_1).$$

If M is a manifold with boundary ∂M, and $\{(\partial M, G, \Phi)\} \in \Gamma(\partial M, G)$ is a given G-action on ∂M, then whether this element is in the image of "∂" is equivalent to ask the question "whether the given action can be equivariantly extended to a G-action on M?"

If $G_1 \overset{i}{\subseteq} G_2$, then a given G_1-action on M, (M, G_1, Φ) can be extended to a G_2-action on M if and only if $\{(M, G_1, \Phi)\} \in \text{Image}\,(h^*)$.

The following are some fundamental but interesting cases.

THEOREM XI. *Let $m \leqq 2n-2$, then the following sequence of semi-groups is exact:*

$$\Gamma(S^{m+1}, SO(n)) \xrightarrow{j} \Gamma(D^{m+1}, SO(n)) \xrightarrow{\partial} \Gamma(S^m, SO(n)) \to 0$$

where j is the operation of removing an equivariant disc around a fixed point in S^{m+1}. I. e. ∂ is onto but is not 1-1 *for $m \geqq n+2$.*

Proof. The onto-ness of ∂ is actually proved in Theorem III and construction 2. In studying all the possible extensions of the orthogonal action into D^{m+1}, we may simply take an action on S^{m+1} and cut off an orthogonal disc around any fixed point. Then the remaining is a D^{m+1} with orthogonal action on the boundary S^m. On the other hand, given any extension of the orthogonal action on S^n into D^{m+1}, we may equivariantly stick back an orthogonal disc to get an action on S^{m+1}. Thus we show the exactness of j, ∂.

THEOREM XII. *Let* $SO(n) \overset{i}{\subseteq} SO(n+1)$, $n+5 \leqq m \leqq 2n$, *then both*

$$\Gamma(S^m, SO(n+1)) \xrightarrow{i^*} \Gamma(S^m, SO(n))$$

and

$$\Gamma(D^m, SO(n+1)) \xrightarrow{i^*} \Gamma(D^m, SO(n))$$

maps everything to the identity, i.e., the element represented by orthogonal action in the two cases repectively.

To speak geometrically, only the orthogonal action can be extended to an action of $SO(n+1)$. and it has in general, infinite many extensions.

Proof. Let $(S^m, SO(n+1), \Phi)$ be a given non-trivial action and X be its orbit space. Then, by Theorem III, X is a compact contractible $(m-n)$-manifold. If we restrict the action to a subgroup $SO(n)$, i.e., we consider the triple $i^*(S^m, SO(n+1), \Phi) = (S^m, SO(n), \Phi \circ (1 \times i))$, we claim that the orbit space $S^m/SO(n)$ will be diffeomorphic to $X \times I$, with corner smoothed. Since $n+5 \leqq m$, we have $m-n \geqq 5$ and by h-cobordism $X \times I \approx D^{m-n+1}$. Now, we apply Theorem III again, we see that $i^*(S^m, SO(n+1), \Phi)$ is orthogonal.

Similar reason for the case of D^m, q. e. d.

Since any $SO(n)$-action on S^m, $n \leqq m \leqq 2n-2$ has at least one fixed point, p. If we remove p from S^m, we get an action on R^m. It follows from Theorem III and V that this defines a well defined map

$$\alpha : \Gamma(S^m, SO(n)) \to \Gamma(R^m, SO(n)).$$

Similarly, if we restrict an $SO(n)$-action on D^m to its interior, we get a map

$$\text{Int}: \Gamma(D^m, SO(n)) \to \Gamma(R^m, SO(n)).$$

It is not difficult to see that we have the following commutative diagram:

$$\begin{tikzcd}
(S^m, SO(n)) \arrow[r, "\alpha"] \arrow[d, "j"] & (R^m, SO(n)) \\
(D^m, SO(n)) \arrow[ur, "\text{Int}"'] &
\end{tikzcd}$$

where j is the same map as defined in Theorem XI.

THEOREM XIII. *Let* $n+3 \leqq m \leqq 2n-2$. *Then*

(i) $\Gamma(S^m, SO(n))$ *and* $\Gamma(D^m, SO(n))$ *are countable but* $\Gamma(R^m, SO(n))$ *is non-countable. Hence both* α *and* Int *are not subjective.*

(ii) Int *is in general not injective.*

Proof.

(i) The fact that $\Gamma(S^m, SO(n))$ and $\Gamma(D^m, SO(n))$ are countable follows immediately from the fact that there are only countable many compact manifolds.

In order to show that $\Gamma(R^m, SO(n))$ is non-countable, we introduce the concept of countable connected chain of discs with action. Let $(D^m, SO(n), \Phi_i)$ be a sequence of D^m with (possibly different) $SO(n)$ action. We let $(M, SO(n), \Psi) = \#\{(D^m, SO(n), \Phi_i)\}$ be the infinite equivariant connected sum with weak topology. If we forget about the boundary of $(M, SO(n), \Psi)$, we get an action on $M^0 = R^m$.

We claim that in this way we may have non-countable action on R^m with their fundamental group of the corresponding sets of fixed points are all different.

(ii) Let $(D^m, SO(n), \Phi_1)$ and $(D^m, DO(n), \Phi_2)$ be two elements of $\Gamma(D^m, SO(n))$ and let (X_1, Y_1) and (X_2, Y_2) be their orbit space respectively. Then $\mathrm{Int}(D^m, SO(n), \Phi_1) = \mathrm{Int}(D^m, SO(n), \Phi_2)$ if and only if $(X_1 - Y_1) \approx (X_2 - Y_2)$.

Hence if the Whitehead group of $\pi_1(\partial Y)$ is not zero and the dimension of ∂Y is high, we may have non-trivial h-cobordism of Y with Y', $Y' \neq Y$. Let $X' = X \cup_Y W$, where W is a given h-cobordism of Y and Y'. Then, surely (X', Y') is another pair of compact contractible manifolds with $(X' - Y') \approx (X - Y)$. q. e. d.

THEOREM XIV. *Let* $n+4 \leqq m \leqq 2n-2$, *then every triple* $(S^m, SO(n), \Phi)$ *can be equivariantly imbedded in* $(S^{n+1}, SO(n), \Phi_0)$, *where* Φ_0 *is the orthogonal action.*

Proof. Let X be the orbit space of $(S^m, SO(n), \Phi)$. By Theorem III, X is a compact contractible manifold of dimension $(m-n+1) \geqq 5$.

Hence $X \times I \approx D^{m-n}$. Again by Theorem III, we may have an $SO(n)$ action on S^{m+1} with $X \times I$ as its orbit space (action is orthogonal, for $X \times I \approx D^{m-n}$). If we consider the inverse image of $X \times \frac{1}{2}$ with restricted action, we have easily the required equivariant imbedding. q. e. d.

BROWN UNIVERSITY
AND PRINCETON UNIVERSITY.

REFERENCES.

[1] J. Adams, "Vector fields on spheres," *Annals of Mathematics*, vol. 75 (1962), pp. 603-632.

[2] R. Bing, "A homeomorphism between the 3-sphere and the sum of two solid horned spheres," *Annals of Mathematics*, vol. 56 (1952), pp. 354-362.

[3] A. Borel, "Impossibilite de fiberer une sphere par un produit de sphere," *Comptes Rendues Academie des Sciences de Paris*, vol. 231 (1950), pp. 943-945.

[4] ———, "Sur la cohomologie des espaces fibres principaux et des espaces homogenes de groupes de Lie compacts," *Annals of Mathematics*, vol. 57 (1953), pp. 115-207.

[5] ———, "Transformation groups with two classes of orbits," *Proceedings of the National Academy of Sciences U. S. A.*, vol. 43 (1957), pp. 983-985.

[6] G. Bredon, "Fixed point sets and orbits of complementary dimension, XV," *Seminar on transformation groups*, Annals of Mathematics Studies 46, Princeton University Press (1961).

[7] ———, "Examples of differentiable group actions" (to appear).

[8] ———, "Transformation groups on spheres with two types of orbits" (to appear).

[9] E. Dynkin, "Maximal subgroups of the classical groups," *Trudy Moskov Mateaticheskikh Obshchesta*, vol. 1 (1952), pp. 39-166. *American Mathematical Society Translations*, Series 2, vol. 6 (1957), pp. 245-378.

[10] M. Hirsch and S. Smale, "On the involutions of the 3 sphere," *American Journal of Mathematics*, vol. 81 (1959), pp. 893-900.

[11] M. Kervaire and J. Milnov, "Groups of homotopy spheres: I," *Annals of Mathematics*, vol. 77 (1963), pp. 504-537.

[12] B. Mazur, "On embeddings of spheres," *Bulletin of the American Mathematical Society*, vol. 65 (1959), pp. 59-65.

[13] D. Montgomery and H. Samelson, "Transformation groups of spheres," *Annals of Mathematics*, vol. 44 (1943), pp. 454-470.

[14] ——— and C. T. Yang, "The existence of a slice," *Annals of Mathematics*, vol. 65 (1957), pp. 108-116.

[15] ——— and L. Zippin, *Topological Transformation groups*, Interscience, 1955.

[16] S. Myers and N. Steenrod, "The group of isometries of a Riemannian manifold," *Annals of Mathematics*, vol. 40 (1939), pp. 400-416.

[17] S. Smale, "On the structure of manifolds," *American Journal of Mathematics*, vol. 84 (1962), pp. 387-399.

[18] P. A. Smith, "Fixed points of periodic transformations Appendix B in Lefschetz," *Algebraic Topology*, New York.

[19] J. Stallings, "The piecewise-linear structure of euclidean space," *Proceedings of the Cambridge Philosophical Society*, vol. 58 (1962), pp. 481-488.

[20] N. Steenrod, *The topology of fibre bundles*, Princeton University Press, 1951.

[21] E. Van Kampen, "On the connection between the fundamental groups of some related spaces," *American Journal of Mathematics*, vol. 55 (1933), pp. 261-267.

ERGODICITY OF FLOWS ON HOMOGENEOUS SPACES.

By Calvin C. Moore.*

Section 1. Let M be a Borel space, and let H be a locally compact group which is separable in the sense of the second axiom of countability. We shall assume that M is analytic (see [20], p. 140; this is merely a regularity condition which will be subsumed later), and we let μ be a σ-finite measure defined on the Borel sets of M. We then suppose that H acts on M as a Borel transformation group (M is a Borel G-space, [19]), denoting the effect of an element $h \in H$ on a point $m \in M$ by $h \cdot m$, and further we assume that the measure μ is quasi-invariant under H. This simply means that $\mu(\sigma) = 0$ if and only if $\mu(h \cdot \sigma) = 0$ for any Borel set σ of M and any $h \in H$. When these conditions are satisfied, it makes sense to speak about the ergodicity of the action of H on M. To be precise, we say that H is ergodic on M if for every Borel set σ of M, $\mu(\sigma \Delta h \cdot \sigma) = 0$ for all $h \in H$ implies that $\mu(\sigma) = 0$ or $\mu(M - \sigma) = 0$. (Δ denotes symmetric difference.) This says that there are no non-trivial elements of the measure algebra of the pair (M, μ) invariant under H. Alternatively, one could define ergodicity as follows: for every Borel set σ of M, $\sigma = h \cdot \sigma$ for all $h \in H$ implies that $\mu(\sigma) = 0$ or $\mu(M - \sigma) = 0$. By a theorem of Mackey ([19], p. 333), these definitions are equivalent. Another characterization is often useful; namely for any Borel function f on M, $f(h \cdot m) = f(m)$ a. e. for $h \in H$ (or $f(h \cdot m) = f(m)$ for all $h \in H$ and $m \in M$) implies that f is almost everywhere equal to a constant function. Furthermore it is clearly enough to consider bounded functions f. The most important case of such actions (or flows as we sometimes call them) from the point of view of ergodic theory is of course when μ is finite and strictly invariant ($\mu(\sigma) = \mu(h \cdot \sigma)$ for $h \in H$ and all Borel sets σ), and when H is the integers or the real line with the usual topology.

Now suppose that H is a subgroup of another group G which also acts on M as a group of Borel automorphisms so that the action of H is obtained by restricting the action of G. We impose firstly the conditions that G be locally compact and separable, that it act as a Borel transformation group on M, and that it preserve the null sets of the measure μ; that is, μ is

---

Received January 21, 1965.

* Research supported in part by National Science Foundation Grant GP-2026.

quasi-invariant under the action of G. Finally we impose the crucial hypothesis that G acts transitively on M. (It is really enough to know that the complement in M of some orbit of G is a μ-null set; the null set may then be deleted without loss of generality.) This is of course a rather severe hypothesis, but many well known flows satisfy the condition (see below). If the conditions above are satisfied, let Γ be the stability group of a point x in M $(\Gamma = \{g; g \in G, g \cdot x = x\})$. Then by [20], p. 148, it follows that Γ is a closed subgroup of G. Moreover it follows from Souslin's theorem (see [20], p. 139) that the natural map of G/Γ onto M by $(g\Gamma) \rightarrow g \cdot x$ is a Borel isomorphism. Thus we may replace M by the homogeneous space G/Γ, and then H acts naturally on G/Γ by left translations $(h \cdot (g\Gamma) = (hg)\Gamma)$. The measure μ on G/Γ obtained by the isomorphism of G/Γ and M is quasi-invariant and hence by [18], p. 104, it is unique up to equivalence. The general question under consideration is to find sufficient conditions for the action of H on G/Γ to be ergodic in terms of the 'position' of H and Γ in G. In the future we shall be considering really only the case when G/Γ admits a finite G-invariant measure. This is a rather severe condition on Γ but it does seem to cover most of the interesting cases.

If in addition to this hypothesis, we also assume that the subgroup H is a one-parameter subgroup of G, many results are known. If for instance G is a finite dimensional vector group, or more generally a connected and simply-connected nilpotent Lie group, simple necessary and sufficient conditions for ergodicity of H on G/Γ are known. (See [1], Chpt. V by L. Green for the nilpotent case.) Weaker results are known if one replaces nilpotence by solvability ([2]). If G is $SL(2, R)$, the group of 2×2 real unimodular matrices, again necessary and sufficient conditions are known. The classical geodesic and horocycle flows are special cases; see [9], [10], [12], [7], [24], and [2] and [21] for some generalizations.

The topic proper of this paper is the study of the case when G is a semi-simple Lie group and H is any subgroup. Theorem 3 in Section 3 below gives sufficient conditions for ergodicity which are almost necessary (see Theorems 4 and 5). The results also show that the same conditions are sufficient to prove weak mixing and strong mixing, when it makes sense for such actions. In the final section we apply these results to prove a theorem which implies as a special case that $SL(2, Z)$, the group of 2×2 integral unimodular matrices, acts ergodically on the plane (Lebesgue measure and the ordinary linear action being understood).

The methods used involve the theory of unitary representations of locally compact groups, and we briefly indicate here the relevance of this subject for the problem at hand. For any $g \in G$, we define a unitary operator $U(g)$ acting on the Hilbert space $L_2(G/\Gamma)$ by $(U(g)f)(m) = f(g^{-1} \cdot m)$, $f \in L_2(G/\Gamma)$. One easily sees that $g \rightarrow U(g)$ is a strongly continuous unitary representation of G—in fact it is the representation of G induced by the trivial one-dimensional representation of Γ ([18]). The crucial fact used in [1], [2], [7], [21] and [24] is that H acts ergodically on G/Γ if and only if the restriction of U to H contains the trivial one-dimensional representation of H at most once. (I.e. the subspace $\{\psi : U(h)\psi = \psi \text{ for all } h \in H\}$ of $L_2(G/\Gamma)$ has dimension at most one, and hence exactly one.) This is clear from our formulation of ergodicity in terms of bounded Borel functions on G/Γ, and the fact that μ is a finite measure. Of course we could have defined the operators $U(h)$ for $h \in H$ without reference to G and so obtain a strongly continuous unitary representation of H. However the key fact is that this representation of H is the restriction to H of the appropriate representation of a larger group G which acts transitively on $M = G/\Gamma$. Our results on ergodic theory are rather easy corollaries of theorems concerned with what happens when one restricts a representation of a semi-simple Lie group G to a subgroup H. The main results are Theorems 1 and 2 which are stated in Section 3. Section 2 is concerned with certain preliminary definitions and simple consequences of them. The rest of the paper after Section 3, except for Section 8, is devoted to the proofs of Theorems 1 and 2.

We wish to thank G. D. Mostow, L. Auslander and T. Sherman for useful discussions about the contents of this paper.

Section 2. We shall now introduce three properties defined in terms of representation theory that a subgroup H of a group G may have. These will be very relevant to our study of ergodicity of flows.

Definition 1. The subgroup H of G has property (E) if for any unitary representation π of G and vector ψ in the Hilbert space of π, the equations $\pi(h)\psi = \psi$ for all $h \in H$ imply that $\pi(g)\psi = \psi$ for all $g \in G$.

In other words, we demand that the trivial one-dimensional representation of H occurs in π restricted to H as a direct summand only as many times as the trivial one-dimensional representation of G occurs in π as a direct summand.

Definition 2. The subgroup H of G has property (WM) if for any (unitary) representation π of G, and *finite dimensional* subspace V of the Hilbert space of π for which $\pi(h)V = V$ for all $h \in H$, we have $\pi(g)v = v$ for all $g \in G$ and $v \in V$.

We note that if H is abelian, then any finite dimensional V as in Definition 2 is the direct sum of one-dimensional H-invariant subspaces. Therefore when H is abelian, it is enough to consider just one-dimensional subspaces V in the above definition, or equivalently, vectors ψ in the Hilbert space of the representation π such that $\pi(h)\psi=\lambda(h)\psi$ for some character λ of H.

In the third definition we want the subgroup H of G to be the real line R, with the usual topology, or the integers Z. If π^* is any (unitary) representation of H, we know by Stone's Theorem (see [16]) that there is a projection valued measure P based on the Borel subsets of $\hat{H}$ (the dual group of H; $\hat{H}=R$ or $\hat{H}=T$, the circle group, in our two cases), such that

$$\pi^*(\phi_t)=\int_{-\infty}^{\infty}\exp(i\lambda t)\,dP_\lambda \text{ if } H=\{\phi_t\}=R$$

or

$$\pi^*(g^n)=\int_0^{2\pi}\exp(in\lambda)\,dP_\lambda \text{ if } H=\{g^n\}=Z.$$

We shall say that π^* has absolutely continuous spectrum if P is absolutely continuous with respect to Lebesgue measure on R or T in the two cases. (That is, $P(\sigma)=0$ if σ is a Lebesgue null set.)

Definition 3. The subgroup H (the line or the integers) has property (SM) if for every representation π of G which does not contain the trivial one-dimensional representation of G, the restriction π^* of π to H has absolutely continuous spectrum. (In the future, when we speak about a subgroup H of G having property (SM) we shall implicitly understand that H is the real line or the integers.)

Our symbolism will be explained below by Proposition 3, but we first want to make a few remarks about weak and strong mixing (WM and SM). Suppose that H acts as a Borel transformation group on a finite measure space (M,μ) leaving μ invariant. We shall say that H is weakly mixing if the 'Cartesian square' of the action is ergodic. That is, let $N=M\times M$ and $\nu=\mu\times\mu$, and let H act on N by $h\cdot(m_1,m_2)=(h\cdot m_1,h\cdot m_2)$. The measure ν is clearly preserved, and if H acts ergodically on N (the Cartesian square of M), we shall say that H is weakly mixing on M. If H is the integers or the line, this is completely equivalent to the usual definition in virtue of the following proposition (see [8]).

PROPOSITION 1. *Let U' be the unitary representation of H on $L_2(M,\mu)$ defined in Section* 1. *(The larger group G is not needed for this.) Then*

H is weakly mixing on M if and only if any finite dimensional subspace V of $L_2(M,\mu)$ invariant under H consists only of constant functions on M.

Proof. This is only a trivial generalization of the corresponding result in [8]. Indeed let H be weakly mixing and let V be a finite dimensional subspace of $L_2(M,\mu)$ invariant under H. If V does not consist just of constants we may assume that V is non-zero and is perpendicular to the constants. Then let $f_1,\cdots,f_n$ be an orthonormal basis of V, and define a function g on $N=M\times M$ by $g(x,y)=\sum_{i=1}^{n}f_i(x)(f_i(y))^*$ $x\in M$, $y\in M$ where $(.)^*$ denotes complex conjugation. Then clearly g is invariant under the Cartesian square flow, is perpendicular to the constant function on N, and is non-zero. Thus the Cartesian square flow is not ergodic.

Conversely suppose that the condition of the proposition holds, and let g be a non-trivial invariant bounded function on N. We may assume without loss of generality that $g(x,y)=g(y,x)^*$ by considering $g(x,y)+g(y,x)^*$ and $i(g(x,y)-g(y,x)^*)$ separately. The formula $(Gf)(x)=\int g(x,y)f(y)d\mu(y)$ defines a bounded linear operator G from $L_2(M,\mu)$ to itself. Clearly for every h in H, we have $GU'(h)=U'(h)G$ using the invariance of g. Thus G is an intertwining operator for the representation U' of H which is self adjoint from the symmetry of g and compact since it is an integral operator. Then for any $\lambda\neq 0$, the null space of $G-\lambda$ is a finite dimensional invariant subspace of U'. Then assuming the condition of the Proposition, we conclude that the nullspace of $G-\lambda$ consists only of constants. Since G is self adjoint, it follows at once that G must be a multiple of the projection onto the constant functions on M. Therefore the original function g must have been a constant almost everywhere, and we are done.

The relevance of condition *WM* above should be clear. We mention now for completeness the definition of strong mixing. If H is the line or the integers, and acts on (M,μ) as above leaving μ invariant, we say that it is strongly mixing if $\lim_{|t|\to\infty}(U'(t)f,g)=(f,1)(1,g)$ for very pair of vectors f and g in $L_2(M,\mu)$ (1 of course is the function constantly equal to one, and $U'(t)$ is as above.) (See [1], [8].)

To return to the main line of development, let us notice that in each of the definitions of properties (*E*), (*WM*), and (*SM*), we universally quantified a representation π of the ambient group G. It makes equally good sense to speak of any one of these three properties relative to a single representation π of G, or more generally relative to some selected family of representations π. A useful instance of this is the following fact.

PROPOSITION 2. *If $H \subset G$ has one of the properties (E), (WM) or (SM) relative to the family of all irreducible representations of G, then H has property (E), (WM) or (SM), respectively.*

Proof. This is merely a formalization of a well known procedure that has been used repeatedly in the past (see [7], Section 2 or [21]). In the first reference, this is done explicitly for (SM). Let us consider (WM); for property (E) the same argument works, even with some simplifications. Let π be a representation of G and V a finite dimensional H-invariant subspace. We can represent π as a direct integral $\int_E \pi^y \, d\mu(y)$ of irreducible representations π^y ([20]). Let $g_1, \cdots, g_n$ be a basis for V. Then in the direct integral decomposition g_i is represented as a function g_i such that $g_i(y)$ is a vector in K^y, the Hilbert space of π^y. Let V^y denote the subspace of K^y spanned by the vectors $\{g_i(y)\}$, $i = 1, \cdots, n$. If $h \in H$, then one sees at once that $\pi^y(g)V^y = V^y$ for almost all y in E. If (h_j) is a countable dense subset of H, then one can extract a μ-null set N from E such that $\pi^y(h_j)V^y = V^y$ for all j and all $y \notin N$. Since π^y is a continuous representation, we see at once that V^y is invariant under all of H for each y not in N. Now we are assuming that (WM) holds for all irreducible representations of G. Thus since V^y is clearly finite dimensional, $\pi^y(g)x = x$ for all $x \in V^y$ if $y \notin N$. Then it is immediate that $\pi(g)f = f$ for any f in the direct integral of the Hilbert spaces K^y such that $f(y) \in V^y$ a.e. In particular this holds for each $g \in V$, and we have proved that (WM) holds for any representation π of G.

To finally justify our previous notation, we state here the consequences of Definitions 1, 2, and 3 for ergodicity of flows on homogeneous spaces.

PROPOSITION 3. *Let G be locally compact, and let Γ be a subgroup such that G/Γ has a finite invariant measure, and let H be a subgroup of G. Then if H has property (E), H acts ergodically on G/Γ. If H has property (WM), then H is weakly mixing on G/Γ, and if H has property (SM), it is strongly mixing on G/Γ.*

Proof. This is again merely a formalization of well known facts. Let U be the natural representation of G on $L_2(G/\Gamma, \mu)$ defined in Section 1. We apply properties E and WM to U and immediately obtain the desired results. (Proposition 1 is needed for weak mixing.) The point here is that the only G-invariant functions in $L_2(G/\Gamma, \mu)$ are the constants. If U^0 is the subrepresentation of U on the orthogonal complement of the constants in $L_2(G/\Gamma, \mu)$, we apply property SM to U^0 to obtain strong mixing. (It is well known that absolute continuity of the spectrum of U^0 restricted to H implies strong mixing (see [1], p. 8).

We shall conclude this section with a comment about the nature of the subgroups H. The point is that we can always assume that H is closed as far as ergodicity questions are concerned.

PROPOSITION 4. *Let G be locally compact, and let Γ be any closed subgroup. Then if H is a subgroup of G, H acts ergodically on G/Γ if and only if $\bar{H}$ (the closure of H) acts ergodically on G/Γ.*

Proof. If G/Γ has a finite G-invariant measure, this is clear from the characterization of ergodicity in Section 1 in terms of unitary representations. However thet proposition is true in general, and we prove it below.

We need only show of course that H is ergodic if $\bar{H}$ is. Thus let σ' be a Borel set in G/Γ invariant under H, and let σ be its inverse image in G. Suppose that we could show that $\mu(\sigma \Delta g \cdot \sigma) = 0$ for every $h \in \bar{H}$ where μ is a left invariant Haar measure on G. (Left Haar measure and right Haar measure have the same null sets.) Then it follows immediately from [18], Lemma 1.3, that $\nu(\sigma' \Delta g \cdot \sigma') = 0$ for $g \in \bar{H}$ where ν is any quasi-invariant measure ([18], p. 103) on G/Γ. Since $\bar{H}$ is assumed to be ergodic with respect to ν, we would be done.

Now to show that $\mu(\sigma \Delta g \cdot \sigma) = 0$ for $g \in \bar{H}$, let I_σ be the characteristic function of σ, viewed as an element of $L_\infty(G)$. Then it is enough to show that

$$\int f(x) I_\sigma(x)\, d\mu(x) = \int f(x) I_{g\cdot\sigma}(x)\, d\mu(x) \quad \text{for } g \in \bar{H},$$

for every $f \in L_1(G, \mu)$ where μ is again left invariant Haar measure on G. Now let $k(g)$ denote the integral on the right hand side of the above as a function on G. From the invariance of μ, we see that

$$k(g) = \int f(g^{-1}x) I_\sigma(x)\, d\mu(x).$$

But $g \rightarrow f(g^{-1}.)$ is a continuous map from G into $L_1(G, \mu)$, and so clearly $k(g)$ is a continuous function of g. We have assumed that $k(h) = k(e)$ for $h \in H$, and so $k(g) = k(e)$ for all $g \in \bar{H}$. This is precisely what we wanted to prove.

Section 3. Suppose now that G is a (connected) semi-simple Lie group with Lie algebra $\mathfrak{g}_0$. Let G^* be the adjoint group of G. Then one knows that G^* is the direct product of its simple factors, $G^* = \prod_{i=1}^{n} G_i^*$, and that $\mathfrak{g}_0$ is the direct sum of the corresponding simple ideals $\mathfrak{g}_0 = \Sigma \oplus (\mathfrak{g}_0)_i$. Now let p denote the projection mapping of G onto G^*, and let p_i denote the projection mapping of G onto G_i^*. If H is a subgroup of G, we shall say that H is *totally non-compact* if $p_i(H)$ has non-compact closure in G_i^* for each i. Finally let us agree to say that G has no compact factors, if no G_i^* is compact. Our main result is the following.

THEOREM 1. *Let G be a (connected) semi-simple Lie group with no compact factors, and let H be a subgroup. Then the following are equivalent:*

1) *H is totally non-compact.*

2) *H has property (E).*

3) *H has property (WM).*

If H is the line viewed as a one-parameter subgroup of G, or if H is the integers viewed as the subgroup generated by a single element of G, then the following is also equivalent:

4) *H has property (SM).*

Of course the non-trivial part of the Theorem ($1 \Rightarrow$ the others) is vacuous if G does happen to have a compact factor since the first condition is *a fortiori* never satisfied in that case. However we shall see that if for instance H is abelian (which covers most of the interesting cases), the theorem remains valid without the hypothesis of no compact factors.

We will indeed prove something which is a little sharper than Theorem 1 (excluding the fourth condition). We note that it is of course enough to replace G by its universal covering group in Theorem 1. Thus assuming G is simply connected, we know that $G = \prod_{i=1}^{n} G_i$ is the direct product of its imple factors where G_i is the universal covering group of G_i^*. The following localization' principle is clearly valid with obvious modifications if G is not imply connected, but it is a bit simpler to state when G is simply connected.

THEOREM 2. *Let π be a unitary representation of G and let H be a subgroup of G. Then let V be a finite dimensional subspace of the Hilbert space of π such that $\pi(h)V = V$ for $h \in H$. Then if $p_i(H)$ has non-compact closure in G_i^*, then $\pi(g)v = v$ for all $g \in G_i$, and $v \in V$.*

The consequences of these results are more or less clear in virtue of Proposition 3 of Section 2.

THEOREM 3. *Let G be a (connected) semi-simple Lie group, and let Γ be a subgroup such that G/Γ has finite G-invariant volume. If H is a totally non-compact subgroup of G, then H is ergodic (and weakly mixing) on G/Γ.*

If H is the line or the integers, and is totally non-compact, then H is strongly mixing on G/Γ.

We shall see shortly that the condition of total non-compactness is at least in a large variety of cases the best possible. We observe that the weak mixing statement is really superfluous since it follows from ergodicity. Indeed if H' is the diagonal subgroup of $H \times H$ in $G \times G$, then H' is clearly totally non-compact in $G \times G$ if H is totally non-compact in G. The Cartesian square (see Section 2) of the flow induced by H on G/Γ is clearly the flow induced by H' on $(G \times G)/\Gamma \times \Gamma \cong G/\Gamma \times G/\Gamma$ which is ergodic if H is totally non-compact. Thus H will be weakly mixing on G/Γ by definition of weak mixing.

In a special case Theorem 2 may be sharpened somewhat as follows:

THEOREM 4. *Let G be simple, non-compact with finite center, and let H be a subgroup. Then if G/Γ has finite G-invariant volume and $\Gamma \neq G$, then H is ergodic on G/Γ if and only if $\bar{H}$ is non-compact.*

Let us now deduce Theorem 3 from Theorem 2 since this leads naturally to the first (and trivial) part of the proof of Theorem 1. Since in Theorem 3, G covers G^*, which is simple, finitely, it is clear that H is totally non-compact if and only if $\bar{H}$ (the closure of H) is non-compact. Thus the 'if' part follows immediately from Theorem 3.

Conversely suppose that H acts ergodically on G/Γ with $\Gamma \neq G$, and that $\bar{H}$ is compact. Clearly we may replace H by $\bar{H}$ and assume that H is compact. But now it is easy to see that a compact group H acts ergodically only if some orbit of H is the complement of a null set in G/Γ. Since any orbit of H is closed, the complement is open in G/Γ, and since the measure μ on G/Γ gives non-zero measure to each non-void open set, we see that H must be transitive. In terms of G, this means precisely that $G = H \cdot \Gamma$. On the other hand, the connected component of Γ is normal in G by a result of Borel [4]. Thus since G is simple, and $\Gamma \neq G$, Γ must be discrete and countable. Now G is of the second category on itself so that $G = \bigcup_{\gamma \in \Gamma} H\gamma$ implies that some $H\gamma$ is not nowhere dense. Clearly this implies that H contains an open set, and in particular a neighborhood of the identity since H is a subgroup. Since G is connected, we have $H = G$ is compact which is contrary to our assumptions.

We can pursue this line of argument somewhat further making use of the sharper information contained in Theorem 2. Let (G_i), $i = 1, \cdots, n$, be a finite set of simple Lie groups with finite centers, and let G be their product. Let Γ be a discrete subgroup of G such that G/Γ has finite measure, such that Γ is skew in G in the sense that $G_i\Gamma$ is dense in G for each i. Such situations can arise in the theory of arithmetic subgroups.

THEOREM 5. *Let G and Γ be as above with Γ discrete and skew in G, and let H be a subgroup of G. Then for H to be ergodic on G/Γ it is necessary and sufficient that $\bar{H}$ be non-compact* $(n \geqq 2)$.

Proof. Since each G_i has finite center, it follows that if $\bar{H}$ is non-compact, then $\overline{p_i(H)}$ is non-compact for some i, say $i = 1$. Now let ψ be a bounded Borel function on G/Γ invariant under H. If U is the natural representation of G on $L_2(G/\Gamma)$, then $U(h)\psi = \psi$ for $h \in H$. Then Theorem 2 implies that $U(g)\psi = \psi$ for any $g \in G_1$. According to the result of Mackey ([18], p. 333) used before, we may modify ψ on a null set in G/Γ such that $\psi(g \cdot m) = \psi(m)$ for all $g \in G_1$ and $m \in G/\Gamma$. We assume this modification done. Now let us denote by f the lift of ψ back to a bounded Borel function on G. Now if $s \in G$, $g \in G_1$ and $\gamma \in \Gamma$, $sgs^{-1} \in G_1$ also since G_1 is normal. Now $f(sg\gamma) = f(sgs^{-1}s\gamma) = f(s\gamma) = f(s)$. Therefore f is invariant under right translation by the group $G_1 \cdot \Gamma$ which by our assumption is dense in G. Now one sees that Proposition 4 above is just as valid for right actions as for left actions, and we apply this to the subgroup $K = G_1 \cdot \Gamma$ of G acting on G itself by right translations. Since G is ergodic on itself, so is K. Therefore the function f above is equal almost everywhere (Haar measure) to a constant. As we have argued before, this implies that the original function ψ on G/Γ is equal almost everywhere to a constant. This completes the proof of the sufficiency of the condition on H.

For the necessity, we assume that $H = \bar{H}$ is compact and acts ergodically on G/Γ. As in Theorem 4, we find that $G = H \cdot \Gamma$, and as Γ is again discrete, we see that $G = H$ is compact. It follows then that Γ is finite, and could not possibly be skew in G. This completes the proof.

Of course one could prove weaker results with fewer assumptions about Γ, but the idea is the same.

We now examine in general the sense in which Theorem 2 is best possible. The following covers a large number of the significant cases.

PROPOSITION 5. *Suppose that H fails to be totally non-compact in G (a connected semi-simple Lie group). Then if G has no compact factors, or if H is abelian, there exists a subgroup Γ such that G/Γ has a finite G-invariant measure, and H is not ergodic on G/Γ.*

Proof. If H fails to be totally non-compact, we may assume that $\overline{p_1(H)} = K$ is compact in $G_1{}^*$. One knows that there exists a discrete subgroup Γ_1 in $G_1{}^*$ such that $G_1{}^*/\Gamma_1$ has a finite $G_1{}^*$-invariant volume (in fact so that $G_1{}^*/\Gamma_1$ is compact, see [5]). Then if $\Gamma = p_1{}^{-1}(\Gamma_1)$, G/Γ has a finite

G-invariant volume. Then if H acts ergodically on G/Γ, the same argument as used in Theorem 4 leads us to the conclusion that $K = G_1^*$. If G has no compact factors, G_1^* is non-compact, and this is impossible. Also if H is abelian, then so is K, and again $K = G_1^*$ is impossible. We remark that this argument could of course be applied under a variety of different conditions on H and G.

Let us now return to Theorem 1. We observe that the implications $4 \Rightarrow 3$ (when (4) makes sense) and $3 \Rightarrow 2$ are trivial. Also we see that $2 \Rightarrow 1$ is trivial for if H is not totally non-compact, Propositions 3 and 5 combined show that H does not have property E. (Note that G is assumed to have no compact factors. Also $2 \Rightarrow 1$ could be done directly without ergodic theory.) Thus we must merely show that (1) implies the other conditions. Clearly $1 \Rightarrow 3$ is taken care of by Theorem 2. We shall treat $1 \Rightarrow 4$ separately for those H for which it makes sense. ((1) $\Rightarrow$ (3) is needed as a technical step in this proof.) The outline of the rest of the paper is as follows: In Section 4 we prove Theorem 2 for one-parameter groups; in Section 5 we prove $1 \Rightarrow 4$ of Theorem 1 for one-parameter groups; in Section 6 we prove both of these results for subgroups generated by a single element; in Section 7 we prove Theorem 2 for an arbitrary subgroup H. Finally in Section 8 we conclude with the applications mentioned in the Introduction.

Section 4. As remarked before, we begin the proof of Theorem 2 with the case when H is a one-parameter subgroup $(\phi_t) = (\exp(tx))$ with $x \in \mathfrak{g}_0$, the Lie algebra of G. We use $(\exp(tx))$ to denote the one-parameter group generated by x. Since H is abelian in this discussion, the same remark as we made after Definition 2 applies; namely it is enough in Theorem 2 to consider a one-dimensional vector space V, or equivalently a vector ψ such that $\pi(\phi_t)\psi = \exp(i\lambda t)\psi$.

Now let $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{h}_0 + \mathfrak{n}_0$ be an Iwasawa decomposition of $\mathfrak{g}_0$ where $\mathfrak{k}_0$ is the Lie algebra of the maximal compact subgroup K^* of G^* and where $\mathfrak{h}_0$ and $\mathfrak{n}_0$ are defined as usual (see [11]). We begin with a very special kind of one-parameter subgroup which has been studied in [21].

LEMMA 4.1. *Theorem 2 is true for any one-parameter subgroup* $H = (\exp(tx))$ *where* $x \in \mathfrak{h}_0$.

Proof. It is clear $\overline{p_i(H)}$ has non-compact closure in G_i^* if and only if $x_i \neq 0$ where $x = \sum x_j$ is the decomposition of x into its components in the simple ideals $(\mathfrak{g}_0)_i$ of $\mathfrak{g}_0$. Now let $\mathfrak{l}_0$ be the subalgebra of $\mathfrak{g}_0$ generated by all vectors y in $\mathfrak{g}_0$ such that $[x, y] = \lambda y$ for $\lambda \neq 0$, and let L be the analytic subgroup of G corresponding to L. Then in [21] (Lemma 7) Mautner essentially proves that if π is any representation of G and ψ a vector in the

Hilbert space of π such that $\pi(\exp tx)\psi = \exp(i\lambda t)\psi$, for all t, then $\pi(g)\psi = \psi$ for all $g \in L$. (Mautner does this for $\lambda = \lambda^*$; for $\lambda \neq \lambda^*$ we refer to the generalization of Mautner's work in [2], Appendix II.)

Thus in order to prove the Lemma, it is enough to show that $\mathfrak{l}_0$ is the sum of those simple ideals $(\mathfrak{g}_0)_i$ of $\mathfrak{g}_0$ such that $x_i \neq 0$. Now if $y = \sum y_i$ satisfies $[x, y] = \lambda y$, clearly $[x_i, y_i] = \lambda y_i$ since $[x_i, y_j] = 0$ if $i \neq j$. Thus it clearly suffices to prove our assertion when $\mathfrak{g}_0$ is simple. That is, if $x \in \mathfrak{h}_0$ and is non-zero, $\mathfrak{l}_0 = \mathfrak{g}_0$. But if $\mathfrak{l}$ and $\mathfrak{g}$ are the complexifications of $\mathfrak{l}_0$ and $\mathfrak{g}_0$, it suffices to prove that $\mathfrak{l} = \mathfrak{g}$. Now $\mathfrak{g}$ is either simple or is the sum of two simple (complex) ideals $\mathfrak{g}^+$ and $\mathfrak{g}^-$ ([11]). If $x \in \mathfrak{g}_0$, then $x = x^+ + x^-$ where both x^+ and x^- are non-zero as is well known. Moreover x, if $\mathfrak{g}$ is simple, and $x^\pm$ if $\mathfrak{g}$ is not simple belong to Cartan subalgebras and every root of $\mathfrak{g}$ is real on these elements. In other words it suffices to consider a complex simple algebra $\mathfrak{g}$ with x, a non-zero real element of a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$; we must show that $\mathfrak{l} = \mathfrak{g}$ where $\mathfrak{l}$ is the algebra spanned by all y with $[x, y] = \lambda y$, $\lambda \neq 0$. This is more or less known, but we include a proof for completeness. We may order the roots of $\mathfrak{g}$ so that x lies in the closure of the positive Weyl chamber of $\mathfrak{h}$. Let $\alpha_1, \cdots, \alpha_n$ be the simple roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$ with this order; then $\alpha_i(x) \geqq 0$ and $\alpha_j(x) > 0$ for some j since $x \neq 0$. Let α_l be any other fundamental root. Since $\mathfrak{g}$ is simple we can find 'chain' $\gamma_1, \cdots, \gamma_k$ of fundamental roots such that $\gamma_1 = \alpha_j$ and $\gamma_k = \alpha_l$, and such that $\sum_{i=1}^{m} \gamma_i$ is a root for any $1 \leqq m \leqq k$. (A similar argument appears in [23].) In particular $\pm\beta_1 = \pm\sum_{i=1}^{k-1} \gamma_i$ and $\pm\beta_2 = \pm\sum_{i=1}^{k} \gamma_i$ are roots. Now since $\gamma_1(x) > 0$ and $\gamma_i(x) \geqq 0$ for $i \geqq 1$, we see that $\beta_i(x) > 0$ for $i = 1, 2$. If X_σ is a root vector for a root σ, then $X_{\pm\beta_i} \in \mathfrak{l}$ for $i = 1, 2$ by the definition of $\mathfrak{l}$. But now $[X_{\beta_2}, X_{-\beta_1}]$ is in $\mathfrak{l}$ and is a non-zero multiple of X_{α_l} since $\beta_2 - \beta_1 = \gamma_k = \alpha_l$. Thus $X_{\alpha_l} \in \mathfrak{l}$ and similarly $X_{-\alpha_l} \in \mathfrak{l}$. But now the vectors $X_{\pm\sigma}$ for fundamental roots σ generate $\mathfrak{g}$, and so $\mathfrak{l} = \mathfrak{g}$.

Remark. The basic idea here is of course Mautner's [21]. In the notation at the beginning of the proof, he showed that $\mathfrak{g}_0 = \mathfrak{m}_0 + \mathfrak{l}_0$ where $\mathfrak{m}_0$ is the centralizer in $\mathfrak{k}_0$ of x, provided each $x_i \neq 0$.

The next step in the proof of Theorem 2 is to prove it for one-parameter subgroups of a special group.

LEMMA 4.2. *Theorem 2 is true for one-parameter subgroups H of the universal covering group G of $SL(2, R)$.*

Proof. We remark here for future reference that once Theorem 2 is proved for some H, it follows automatically for any subgroup gHg^{-1} conjugate to H in G. In the present case there are only three conjugacy classes of one-parameter subgroups. Those of elliptic type ($p(H)$ compact in $SL^*(2,R)$) fail to satisfy the hypotheses. Those of hyperbolic type ($p(H)$ conjugate to $(\exp(tx))$, $x = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$) in $SL^*(2,R)$) are taken care of by [21] or [2] or Lemma 4.1. Those of parabolic type ($p(H)$ conjugate to $(\exp(tx))$, $x = \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}$ in $SL^*(2,R)$) remain. Now Parasyuk in [24] has proved what we want for $SL(2,R)$ in place of its covering group G, by checking it for each irreducible representation π of $SL(2,R)$. One could presumably do the same for G. However this is not necessary since Lemma 2 of [27] gives the non-trivial part of what we need. (We are indebted to Sherman for showing us a prepublication copy of his work.)

His result is that if π is any representation of G and $\pi(\exp tx)\psi = \psi$ for some ψ, then $\pi(g)\psi = \psi$ for $g \in G$. It remains to consider the case when $\pi(\exp tx)\psi = \exp(i\lambda t)\psi$ for some ψ and some $\lambda \neq 0$. Now let y be an element of the Lie algebra of G such that $[y,x] = -x$ (such clearly exist). Then if $\psi_a = \pi(\exp ty)\psi$, one sees at once by a simple calculation that $\pi(\exp tx)\psi_a = \exp(ie^a\lambda t)\psi_a$. Thus ψ_a is a simultaneous eigenvector for the group $(\pi(\exp tx))$ of operators and since $\lambda \neq 0$, the corresponding characters $(t \rightarrow \exp(ie^a\lambda t))$ are distinct. It follows that for $a \neq b$, $(\psi_a, \psi_b) = 0$. This contradicts the separability of the Hilbert space of π unless $\psi = 0$, and in this case $\pi(g)\psi = \psi$ for $g \in G$, and we are done.

The next lemma proves Theorem 2 for one-parameter subgroups of the opposite extreme from Lemma 4.1.

LEMMA 4.3. *Theorem 2 is valid for any G, and for any one-parameter subgroup $H = (\exp(tx))$ where $x \in \mathfrak{g}_0$ is nilpotent (i.e. $ad(x)^n = 0$ on $\mathfrak{g}_0$).*

Proof. By the Jacobson-Morosov lemma ([22], [13], [15]) we may find a three-dimensional simple non-compact subalgebra $\mathfrak{b}_0$ of $\mathfrak{g}_0$ which contains x. Now $\mathfrak{b}_0$ is the Lie algebra of $SL(2,R)$, and we may clearly choose coordinates here so that $(di)\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix} = x$ where di is the injection of $\mathfrak{b}_0$ into $\mathfrak{g}_0$. Now let i be the corresponding homomorphism of the universal covering group G_1 of $SL(2,R)$ into G.

Now if π is a representation of G such that $\pi(\exp(tx))\psi = \exp(i\lambda t)\psi$ for some vector ψ, let $\pi^* = \pi \circ i$ be the corresponding representation of G_1. Then $\pi^*(\phi_t)\psi = \exp(i\lambda t)\psi$ where ϕ_t is a parabolic type one-parameter subgroup of G_1. It follows by 4.2 that $\pi^*(g)\psi = \psi$ for $g \in G_1$. In particular if $y = (di)\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$, we see that $\pi(\exp(ty))\psi = \psi$ for all t. Note that $[y,x] = x$.

But now $ad(y)$ on $\mathfrak{g}_0$ is diagonalizable with real eigenvalues. The one-parameter group $(\exp^*(ty))$ in the adjoint group G^* of G is semi-simple and diagonalizable over the reals. It may then be embedded in a maximal (algebraic) torus T diagonalizable over the reals ([25]). But then one knows that T is conjugate to $\exp^*(\mathfrak{h}_0)$ in G^* where $\mathfrak{h}_0$ is the Euclidean factor in the Iwasawa decomposition $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{h}_0 + \mathfrak{n}_0$ of $\mathfrak{g}_0$. Thus we can change this decomposition so that T is $\exp^*(\mathfrak{h}_0)$. Then we find that $y \in \mathfrak{h}_0$ and Lemma 4.1 is applicable to $(\exp(ty))$. To be precise, if $y = \sum y_i$ is the decomposition of y in terms of components in the simple factors of $\mathfrak{g}_0$, then $\pi(g)\psi = \psi$ for $g \in G_i$ if $y_i \neq 0$.

To go back to our original one-parameter subgroup $H = (\exp(tx))$, we want to know that $\pi(g)\psi = \psi$ for $g \in G_i$ provided $\overline{p_i(H)}$ is non-compact. But clearly $p_i(H)$ is non-compact if and only if $x_i \neq 0$ where x_i is the component of x in $(\mathfrak{g}_0)_i$. But now it follows from the relation $[y, x] = x$ that $[y_i, x_i] = x_i$, and so if $x_i \neq 0$, then $y_i \neq 0$. Once $y_i \neq 0$, we have the desired result using the preceding paragraph.

Now let $H = (\exp(tx))$ be a general one-parameter subgroup. Then one knows that $x = x' + x''$ where $ad(x')$ is semi-simple and $ad(x'')$ is nilpotent, and $[x', x''] = 0$ (see [3]). This decomposition is unique. Moreover, if $x = \sum x_i$ is written as the sum of its components in the simple ideals, then $x' = \sum x_i'$ and $x'' = \sum x_i''$ where x_i' and x_i'' are the semi-simple and nilpotent parts of x_i. The eigenvalues of $ad(x)$ and $ad(x')$ (and of $ad(x_i)$ and $ad(x_i')$) are the same.

LEMMA 4.4. *If* $H = (\exp(tx))$, *then* $\overline{p_i(H)}$ *is non-compact if and only if either* $ad(x_i)$ *has an eigenvalue with non-vanishing real part or* $x_i'' \neq 0$.

Proof. This is absolutely clear.

LEMMA 4.5. *If* $H = (\exp(tx))$ *is a one-parameter subgroup of the semi-simple group* G, *and if* $ad(x_i)$ *has an eigenvalue with non-vanishing real part, then Theorem 2 holds for this* i.

Proof. We may suppose that $ad(x_i)$ has an eigenvalue with positive real part. Now let y be an eigenvector of $ad(x_i)$ in $\mathfrak{g}_i$ (the complexification of $(\mathfrak{g}_0)_i$) such that the eigenvalue λ has maximal real part. Thus $[x_i, y] = \lambda y$ and if σ is the conjugate linear involution of $\mathfrak{g}_i$ fixing $(\mathfrak{g}_0)_i$ then $[x_i, \sigma(y)] = \lambda^* \sigma(y)$ where λ^* is the complex conjugate of λ. Then one sees at once that $[y, \sigma(y)]$ is an eigenvector of $ad(x_i)$ with eigenvalue $\lambda + \lambda^*$. Now since λ had maximal (positive) real part, $[y, \sigma(y)] = 0$.

It is known that $[z, y] = \mu y$ for $\mu \neq 0$ in a Lie algebra implies that $ad(y)$ is nilpotent ([14], p. 44). Thus $ad(y)$ and $ad(\sigma(y))$ are nilpotent, and since $[y, \sigma(y)] = 0$, $ad(z)$ for any z in the linear space $\mathfrak{n}$ spanned by y

and $\sigma(y)$ is nilpotent. We agree to choose y such that $y=\sigma(y)$ if λ is real. Now $\sigma(\mathfrak{n})=\mathfrak{n}$, and let $\mathfrak{n}_0=\mathfrak{n}\cap\mathfrak{g}_0$, and let $\mathfrak{s}_0$ be the two- or three-dimensional solvable algebra spanned by x and $\mathfrak{n}_0$. The analytic subgroup S of G corresponding to $\mathfrak{s}_0$ satisfies the conditions of the version of Mautner's lemma given in [2], Appendix II.

Indeed if π is a representation of G such that $\pi(\exp tx)\psi=\exp(i\lambda t)\psi$ for some ψ in the Hilbert space of π, the result quoted above tells us that $\pi(s)\psi=\psi$ for all $s\in S$. In particular $\pi(\exp tz)\psi=\psi$ for some non-zero z in $\mathfrak{n}_0$. But $ad(z)$ is nilpotent, and Lemma 4.2 tells us that $\pi(g)\psi=\psi$ for all $g\in G_i$ where G_i is the analytic subgroup of the simple factor $(\mathfrak{g}_0)_i$ under consideration. This completes the proof of the lemma.

Remark. For conversational convenience let us call a two- or three-dimensional solvable Lie algebra $\mathfrak{s}_0$ or its corresponding simply connected group S isomorphic to one of types of algebras, constructed in this lemma, a standard algebra or group. We will denote by $\mathfrak{n}_0$ and N the Lie algebra and subgroup defined above. We note that $\mathfrak{n}_0$ is the derived algebra and the nilradical of $\mathfrak{s}_0$. The above proof shows that if $x\in\mathfrak{g}_0$, and $ad(x)$ has an eigenvalue which is not purely imaginary, then x is in some standard subalgebra. Also if $ad(x)$ is nilpotent, then the Jacobson-Morosov lemma above shows that x is in some standard (two-dimensional) subalgebra. This will be used later.

We now complete the proof of Theorem 4 for one-parameter groups by treating the remaining case in Lemma 4.5.

LEMMA 4.6. *Let $H=(\exp(tx))$ and suppose $x_i''\neq 0$ for some i. If π is a representation of G and ψ a vector in the Hilbert space of π such that $\pi(\exp tx)\psi=\exp(i\lambda t)\psi$, then $\pi(g)\psi=\psi$ for $g\in G_i$.*

Proof. As we have remarked before, it suffices to take π irreducible which we now do. If F is the set of indices i such that $ad(x_i)$ has an eigenvalue with non-zero real part, then we already know that $\pi(g)\psi=\psi$ for $g\in G_i$, $i\in F$. Now the set of vectors β with $\pi(g)\beta=\beta$ for $g\in G_i$, $i\in F$ is clearly a closed invariant subspace V of the representation space. It is non-zero since we may assume that the vector ψ in the statement of the lemma is non-zero. Since π is irreducible, V is the entire Hilbert space of π.

Therefore G_i, $i\in F$ acts completely trivially under π, and we may as well replace H by its projection into the remaining factors, and drop the G_i, $i\in F$. In other words we may assume that $ad(x)$ has only purely imaginary eigenvalues. Then $x=x'+x''$, and let $H'=(\exp(tx'))$ be the one-parameter subgroup generated by x'. Now $p(H')$, the projection of H' into the adjoint group, has compact closure A^*. Then A^* is compact connected and abelian.

Let A be the connected component of the identity of $p^{-1}(A)$; then $H' \subset A$, and A centralizes H. Moreover if Z is the center of G, we see at once that $A^* \cong A/(A \cap Z)$ by means of the projection p.

Since π is irreducible, $\pi(z) = \sigma(z) \cdot 1$ for $z \in Z$, where σ is some character of Z. Now let W be the subspace of all vectors in the Hilbert space of π such that $\pi(\exp tx)\psi = \exp(i\lambda t)\psi$. Then W is closed, and since A centralizes H, W is invariant under $\pi(A)$. Let π^* denote the representation of A on W so defined. Then $\pi^*(a) = \sigma(a) \cdot 1$ for $a \in A \cap Z$ by the above. Since $A/A \cap Z$ is compact, it follows that π^* decomposes as a discrete direct sum of multiples of irreducible representations of A. To be precise, $W = \Sigma \oplus W_\nu (\nu \in \hat{A})$ where $W_\nu = \{\beta : \pi^*(a)\beta = \nu(a)\beta\}$. The sum need only be taken over the (denumerable) subset of $\hat{A}$ consisting of those ν for which $\nu(a) = \sigma(a)$ for $a \in A \cap Z$. If $\beta \in W_\nu$, we see that $\pi(\exp tx'')\beta = \exp(ist)\beta$ for some fixed s. We know by Lemma 4.3 then that $\pi(g)\beta = \beta$ for $g \in G_i$ if $x_i'' \neq 0$. Therefore $\pi(g)\alpha = \alpha$ for every $\alpha \in W$, and in particular for ψ, if $g \in G_i$ and $x_i'' \neq 0$.

Theorem 2 is now completely proved for all one-parameter subgroups, and hence the implication $1 \Rightarrow 3$ of Theorem 1 is proved for one-parameter subgroups. At this point it will be convenient to take care of $1 \Rightarrow 4$ of Theorem 1 for one-parameter subgroups before we treat Theorem 2 for general subgroups H.

Section 5. Suppose that $\mathfrak{g}_0$ is a semi-simple Lie algebra and G is its corresponding simply connected group. Let $\mathfrak{s}_0$ be a standard solvable subalgebra of $\mathfrak{g}_0$ (see the remark after Lemma 4.5) with nilradical $\mathfrak{n}_0$. Then we may view the simply connected Lie group S corresponding to $\mathfrak{s}_0$ as a subgroup of G; similarly for the simply connected group N. Then N is a vector space of dimension one or two as is $\hat{N}$, the dual group of N. We choose a vector y in $\mathfrak{s}_0$ not in $\mathfrak{n}_0$, such that the real part of the eigenvalues of $ad(y)$ on $\mathfrak{n}_0$ are one.

We coordinatize N when it has dimension two so that $ad^*(\exp(ty))$ ($*$ denotes transposed action on $\hat{N}$) acts by the matrix

$$e^t\begin{pmatrix} \cos \alpha t & \sin \alpha t \\ -\sin \alpha t & \cos \alpha t \end{pmatrix}$$

where α is the magnitude of the imaginary part of the eigenvalues of $ad(y)$. If $\hat{N}$ is one-dimensional, $ad^*(\exp ty)$ acts by multiplication by e^t. We use $|\cdot|$ to denote the Euclidean length of a vector using these coordinates.

The following lemma will take care of the crux of the matter of proving $1 \Rightarrow 4$ in Theorem 1.

LEMMA 5.1. *Suppose that* $\mathfrak{s}_0$ *is a standard subalgebra of* $\mathfrak{g}_0$ *and let* $z \in \mathfrak{s}_0$ $z \neq 0$. *If* π *is any unitary representation of the simply connected* *semi-simple group G such that no simple factor G_i leaves fixed a non-zero vector in the Hilbert space of π, then $(\pi(\exp(tz)))$ has absolutely continuous spectrum.*

Proof. We use the notation above. Clearly it suffices to prove the lemma for a representative of each conjugacy class of one-parameter subgroups in $\mathfrak{s}_0$. Also we may change scale arbitrarily (i.e. replace z by sz, $s \neq 0$). Now if z is not in $\mathfrak{n}_0$ we may clearly take $z = x$ above, by using these normalizations. (This is not really essential, but makes notation simpler.) Now let π^* denote the restriction of π to N. Since N is abelian, Stone's theorem [16] yields a canonical projection valued measure P based on the Borel sets of $\hat{N}$. We shall define a new projection valued measure Q based on the Borel subsets of line R. Indeed if $\sigma \subset R$, let $\sigma^* = \{x : x \in \hat{N}, \log|x| \in \sigma\}$ where $|\cdot|$ is the norm on $\hat{N}$ defined above. Then let $Q(\sigma) = P(\sigma^*)$ noting that σ^* is a Borel subset of $\hat{N}$ if σ is a Borel subset of line. (This is really just a logarithmic change of variable; the same device is used by Sherman in [27] for the same purpose.) We must now note that Q is really a projection valued measure, the only difficulty being that $Q(R) = 1$. This will be true if $P(\{0\}) = 0$ where 0 is the neutral element of $\hat{N}$. But $P(\{0\})$ is just the projection onto those vector ψ such that $\pi(g)\psi = \psi$ for $g \in N$. Now we know that every element of $\mathfrak{n}_0$ is nilpotent in $\mathfrak{g}_0$. Thus there exists a simple factor G_i, such that $(p_i(\exp(tn)))$ has non-compact closure for some $n \neq 0$ in $\mathfrak{n}$. Then $\pi(\exp(tn))\psi = \psi$ implies by Theorem 2 that $\pi(g)\psi = \psi$ for $g \in G_i$. Our hypotheses on π now imply that $\psi = 0$. Thus $P(\{0\}) = 0$.

But now the action of $ad^*(\exp tx)$ on $\hat{N}$ is known; to be precise, if σ is a Borel set of R, and σ^* is as above, then $ad^*(\exp(tx)) \cdot \sigma^* = (\sigma + t)^*$ where $\sigma + t = \{x + t ; x \in \sigma\}$. A trivial calculation shows that $Q(t + \sigma) = \pi(\exp(-tx))Q(\sigma)\pi(\exp(tx))$. Then if we define a one-parameter continuous group (V_s) of unitary operators by $V_s = \int \exp(i\lambda s)\, dQ_\lambda$, we find that $V_s\pi(\exp tx) = \exp(its)\pi(\exp tx)V_s$. These are exactly the Weyl commutation relations, and since they have a standard model ([17]), we find at once from this standard model that $(\pi(\exp tx))$ and (V_s) have absolutely continuous spectrum.

Thus we have proved the Lemma for $(\pi(\exp tz))$ if $z \notin \mathfrak{n}_0$. Now suppose that $z \in \mathfrak{n}_0$. Then $ad(z)$ is nilpotent and by our previous remarks z belongs to the nilradical $\mathfrak{n}_0$ of some new two-dimensional standard subalgebra $\mathfrak{s}_0$ of $\mathfrak{g}_0$.

We change the meaning of the symbols in the previous paragraph and let them apply to our new $\mathfrak{s}_0$. Now N and $\hat{N}$ are the line and the projection valued measure P on the Borel set of N is the spectral resolution for the one-parameter group $\pi(\exp(tz))$. We know that $P(\{0\}) = 0$, and we must show that if τ' is a Lebesgue null set not containing the origin, then $P(\tau') = 0$.

Now $\tau = \tau' \cup (-\tau')$ is a Lebesgue null set not containing the origin and τ is of the form σ^* for some Borel set σ of R. Indeed, $\sigma = \log(\tau)$ so that σ is a null set also. Therefore, since Q is absolutely continuous, $Q(\sigma) = P(\tau) = 0$. Since $\tau' \subset \tau$, we have $P(\tau') = 0$ also as desired.

The main part of the next lemma takes care of the same exceptional case that we had to do separately in the last section.

LEMMA 5.2. *If G is semi-simple, then any totally non-compact one-parameter subgroup has property SM.*

Proof. Let $H = (\exp(tx))$, $x \in \mathfrak{g}_0$. If $ad(x)$ has an eigenvalue with non-zero real part, or if $ad(x)$ is nilpotent, then x can be embedded in a standard solvable subalgebra $\mathfrak{s}_0$ of $\mathfrak{g}_0$ by the remark following Lemma 4.5. The only other way that H can be totally non-compact is for $x_i'' \neq 0$ for all i and for x to have all eigenvalues purely imaginary where $x = x' + x''$ is the decomposition of x into its semi-simple and nilpotent parts. We note finally that it is no loss of generality to take G to be simply connected; then $G = \prod_{i=1}^{n} G_i$ with G_i simple.

First let us consider an irreducible representation π of G with the property of Lemma 5.1; namely that no simple factor of G_i has a non-zero invariant vector. In the first case when $ad(x)$ has an eigenvalue with non-zero real part, Lemma 5.1 gives us the result. Then we consider the remaining case above. The argument in Lemma 4.6 may be applied to $\exp(tx')$, and shows that $(\pi(\exp(tx')))$ has pure discrete spectrum. Also we already know that $(\pi(\exp(tx'')))$ has absolutely continuous spectrum. Then if $d\pi(z)$ denotes the infinitesimal generator of $(\pi(\exp(z)))$ for $z \in \mathfrak{g}_0$, we have $d\pi(x) = d\pi(x') + d\pi(x'')$, and the final two operators commute in the sense that their spectral resolutions commute. One has pure point spectrum and the other has absolutely continuous spectrum. It is virtually obvious and easy to prove that the sum of two such operators has absolutely continuous spectrum.

To summarize, we have shown that $H = (\exp(tx))$ has property (SM) for every irreducible representation π of G such that no simple factor G_i of G leaves fixed a non-zero vector in the Hilbert space of π. We claim that this is enough. Indeed to establish (SM), it is enough to consider an arbitrary irreducible representation π. Now let F be the set of indices i such that G_i has a non-zero fixed vector in the space of π. If $i \in F$, let $V_i = \{\psi : \pi(g)\psi = \psi, g \in G_i\}$. Then V_i is a closed subspace of the Hilbert space of π, and is invariant under G since G_i is normal in G. Since π is irreducible, V_i is everything. In other words $G_F = \prod_{i \in F} G_i$ acts completely trivially under π. Let G^1 be the product of the remaining factors and $H^1 = \exp(tx^1)$ the projection of H in G^1 and π^1 the restriction of π to G^1. (Note that $x = x^1 + x^2$ with x^2 in $\sum_{i \in F} (\mathfrak{g}_0)_i$ and x^1 in the sum of the remaining factors.) Then $\pi(\exp tx) = \pi^1(\exp tx^1)$, and we can replace G, H and π by G^1, H^1 and π^1. Now our side condition that no simple factor of G^1 has any non-zero invariant vectors under π^1 is satisfied. Finally note that H^1 is still totally non-compact in G^1. This completes the proof of Lemma 5.2.

For one-parameter groups of the type $(\exp tx)$ with x in $\mathfrak{h}_0$ (see Lemma 4.1) this extends known results in [7] and [21]. In this regard we remark that Lemma 14 in [21] is false as it stands.

Section 6. We shall now consider a discrete group H generated by a single element $g \in G$. We want to prove Theorem 2 and $1 \Rightarrow 4$ of Theorem 1 for this class of subgroups of G.

LEMMA 6.1. *Theorem 2 holds for any subgroup H generated by a single element g in G.*

Proof. Let $g^* = p(g)$ be the image of g in G^*, the adjoint group of G. If T is the algebraic hull of the group $p(H)$ in the linear algebraic group G^*, then T has a finite number of components. Thus for some n, $(g^*)^n \in T_0$, the component of the identity of T. Since T_0 is a connected abelian Lie group, there is a one-parameter subgroup $(\exp^*(tx))$ in G^* such that $\exp^*(x) = (g^*)^n$. It follows then that $\exp(x) = z \cdot g^n$ for some z in the center Z of G.

For the proof of Theorem 2, we need only consider an irreducible representation π of G and a vector ψ in the Hilbert space of π such that $\pi(g)\psi = \exp(i\alpha)\psi$. Since π is irreducible, it is scalar on Z and we find that $\pi(\exp(x))\psi = \exp(i\beta)\psi$ for suitable β. Now let V be the set of all vectors ϕ in Hilbert space of π which satisfy $\pi(\exp(x))\phi = \exp(i\beta)\phi$. Clearly V is closed and invariant under $\pi(\exp tx)$ for all t. Now by the same argument as used in Lemma 4.6, $V = \sum V_\lambda$ where $V_\lambda = \{\phi : \phi \in V, \pi(\exp tx)\phi = \exp(i\lambda t)\phi\}$ where we sum over only those λ such that $\lambda = \beta \bmod (2\pi)$. Suppose that $\phi \in V_\lambda$; then by Theorem 2 for one-parameter groups, $\pi(g)\phi = \phi$ for all $g \in G_i$ if $p_i((\exp(tx)))$ has non-compact closure. Therefore the same holds for all $\phi \in V$, and in particular for our original vector ψ. But now suppose that $(p_i(g))$ has non-compact closure in G_i^*, then clearly so does $p_i((\exp(tx)))$ since $p_i(g^n) = p_i(\exp(x))$. Therefore we have the desired result.

We now come to $1 \Rightarrow 4$ of Theorem 1 for subgroups H generated by a single element of G.

LEMMA 6.2. *If $H=\{g^k\}$, $-\infty<k<\infty$ is totally non-compact in G, then H has property SM.*

Proof. As in the previous lemma, we can find an integer n, an element z of Z, the center of G, and a totally non-compact one-parameter subgroup $K=(\exp(tx))$ such that $g^n=z\exp(x)$. As before we know that it is enough to consider a non-trivial irreducible representation π of G. Then $\pi(z)=\exp(i\alpha)$. We write the spectral representations of $\pi(g^k)$ and $\pi(\exp tx)$ as

$$\pi(g^k)=\int_0^{2\pi}\exp(ik\lambda)\,dP_\lambda$$

$$\pi(\exp tx)=\int_{-\infty}^{\infty}\exp(it\lambda)\,dQ_\lambda$$

for projection valued measures P and Q. We must show that $P(\sigma)=0$ for Borel null sets σ. Now suppose that σ is a Borel null subset of $[0,2\pi)$, and then let $\sigma^*=\{x\colon x\in R, x+\alpha=y(\operatorname{mod}2\pi/n)$ for some $y\in\sigma\}$. Then it is easy to see that σ^* is a Borel subset of R of zero Lebesgue measure. The equation $\pi(g^n)=\exp(i\alpha)\pi(\exp(x))$ then implies that $P(\sigma)=Q(\sigma^*)$. Since K is totally non-compact $Q(\sigma^*)=0$, and hence $P(\sigma)=0$ as desired.

Section 7. We now come to the proof of Theorem 2 for general subgroups H of G. As remarked before, this will complete the proof of Theorem 1 since we have already taken care of the implication $1\Rightarrow 4$ when it has meaning. Thus as before, we consider an irreducible representation π of G (which is simply connected) and a finite dimensional non-zero subspace V of the Hilbert space such that $\pi(h)V=V$ for $h\in H$. Then if $p_i(H)$ for some i has non-compact closure, we want to know that $\pi(g)v=v$ for $g\in G_i$ and $v\in V$. Now we may replace H by $H\cdot Z$ where Z is the center of G since $\pi(z)=\lambda(z)\cdot 1$ for $z\in Z$. Also we may replace H by its closure $\bar{H}$ and assume without loss of generality that H contains Z and is closed. Under these conditions $p(H)\subset G^*$ is closed.

Now let F be the set of indices j such that there exists a $g\in H$ such that $\{p_j(g^n)\}$, $-\infty<n<\infty$, has non-compact closure in G_j^*. Then by Lemma 6.1, $\pi(g)v=v$ for all $g\in G_j$, $v\in V$ if $j\in F$. If our original index i is in F, we are done, so we suppose that it is not. Now the set of vectors x such that $\pi(g)x=x$ for all $g\in G_j$, $j\in F$ is a non-zero closed invariant subspace of the Hilbert space of π, and since π is irreducible, it is the total space. (It is non-zero since it contains V.) Thus the G_j for $j\in F$ act trivially under π so that we can replace H by $H\cdot(\prod_{j\in F}G_j)$ and then factor out by $\prod_{j\in F}G_j$. Equivalently, we replace G by the product of the G_j for $j\notin F$, and H by its projection onto this group. We change our notation. Our new

H may not be closed, so we close it, and it still contains the center of G. We can repeat the above process on H, and possibly remove some more factors, and so on until the process stabilizes.

We eventually arrive with a group G, a closed subgroup H containing the center such that for any $h \in H$, and simple factor G_j, $\{p_j(h^n)\}$ has compact closure in G_j^*. This means simply that $\{p(h^n)\}$ has non-compact closure in G^*. We may of course assume that our original index i such that $\overline{p_i(H)}$ is non-compact still lives in G; otherwise we would have been done. Then in particular $p(H)$ is non-compact. If we look at $H^* = p(H)$ in G^*, we have a non-compact subgroup H^* such that for any element $h^* \in H$, the set $((h^*)^n)$ has compact closure. To complete the proof it will be enough to show that this state of affairs is impossible. The crucial fact is that G^* is a matrix group.

LEMMA 7.1. *Let H be a closed subgroup of $GL(n,R)$ such that the set $\{h^n\}$ has compact closure for each $h \in H$. Then H is compact.*

Proof. Let H_0 be the connected component of the identity in H. Then H_0 is closed in $GL(n,R)$ and is a connected Lie group. If it is non-compact, it will certainly have a closed one-parameter subgroup (ϕ_t) which is non-compact. Then the set of elements $\{\phi_n\} = \{(\phi_1)^n\}$ do not have compact closure. Therefore H_0 is compact.

Since H is itself a Lie group, H_0 is open in H. If $h \in H$, then h is contained in a compact group $K \subset H$, and it follows that some subsequence of (h^n) must converge to the identity element of K. Since $K \cap H_0$ is open in K, we find that $h^k \in H_0$ for some k. Therefore H/H_0 is a torsion group. Also since any $h \in H$ is contained in a compact subgroup, H is in fact contained in $SL(n,R)$.

Now since H_0 is compact, we can choose coordinates such that $H_0 \subset SO(n,R)$, the orthogonal group. Now let A be the algebraic hull of H, and let A^0 be the algebraic hull of H_0. Since $H_0 \subset SO(n,R)$, and the latter group is algebraic, A^0 is a compact group. Since H_0 is normal in H, it is normal in its algebraic hull, since the normalizer of H_0 is an algebraic subgroup. Furthermore A^0, the algebraic hull of H_0 is normal in A.

Now let $B = H \cdot A^0$; this is a closed subgroup of A since A^0 is compact and normal in A, and H is closed. Now $B \supset A^0$ and B/A^0 is a torsion group since H/H_0 was and $A^0 \supset H_0$. To prove the Lemma it is enough to know that B is compact or what is the same, that B/A^0 is compact. Now the crucial point is that A/A^0 is also a linear algebraic group since A is linear and A^0 is normal in A ([3], p. 40). Thus the Lemma comes down to a special case of itself, namely when the matrix group H (here B/A^0) is a torsion group.

If in the following lemma H is finitely generated, this is just a classical result of Burnside (see [26] and [5] for generalizations of this).

LEMMA 7.2. *Let H be a closed subgroup of $GL(n,R)$ such that every element has finite order. Then H is finite.*

Proof. Observe that the connected component of H is trivial since any non-trivial connected Lie group has elements of infinite order. Thus H is discrete and hence countable. If F is a finitely generated subgroup of H, then the result of Burnside mentioned above applies, and tells us that F is finite. Thus H can be exhausted by an ascending chain F_n of finite groups.

For any subgroup A of H, let $M(A)$ denote the algebra of finite linear combinations with complex coefficients of elements of A (the group algebra of A). Let π denote the natural complex representation of H on V^n, a complex n-dimensional vector space. We extend π in the well known way to $M(H)$ as a homomorphism of $M(H)$ into the endomorphism ring of V^n.

Since $\pi(M(H))$ is necessarily finite dimensional, we see at once that $\pi(M(H))=\pi(M(F_k))$ for some finite subgroup F_k. In particular, a subspace W of V^n is invariant under the operators $\pi(H)$ if and only if it is invariant under the operators $\pi(F_k)$ for this k. Thus the representation π of H is completely reducible since π restricted to F_k is. Let $V^n=\Sigma\oplus V_i$ be represented as a direct sum of invariant irreducible subspaces for H (and F_k). Since F_k is finite, there is a positive definite Hermitian form f_i on V_i which is invariant under F_k. Moreover f_i is unique up to positive scalars. Now if $F_l\supset F_k$ and is finite, then F_l also leaves invariant some positive definite Hermitian form on V_i. By the unicity, F_l must leave f_i invariant. Since the F_l exhaust H, we see that H itself must leave f_i invariant. Then we can assemble the various f_i to form a positive definite Hermitian form for V by making the subspace V_1 mutually clearly orthogonal. The form f is invariant under H so that H is contained in a unitary group on V^n. Since H is closed, it follows that H is compact, and since as we have already noticed, H is discrete, H must necessarily be finite.

Now the results announced in Section 3 are completely proved.

Section 8. We should like to preface this section by saying that the ideas and main results of this section were also found independently and essentially simultaneously by G. W. Mackey.

Up to now we have been considering the action of a subgroup H of G on the homogeneous space G/Γ of cosets of the form $g\cdot\Gamma$. We noted in Section 2 (Proposition 4) that from the point of view of ergodicity questions, it was no loss of generality to take H closed. Let us change notation, and consider a locally compact (separable) group G with two closed subgroups H_1 and H_2. We make no assumptions whatever about the existence of invariant measures on the homogeneous spaces. As before H_1 acts on G/H_2 by $h_1 \cdot (gH_2) = (h_1 g)\Gamma$, and also we see that H_2 acts on $H_1\backslash G$ by $h_2 \cdot (H_1 g) = H_1(gh_2^{-1})$. (We use $H_1\backslash G$ to denote the family of cosets of the form $H_1 g$ of the opposite parity. Since G/H_2 and $H_1\backslash G$ always have quasi-invariant measures, it still makes sense to talk about ergodicity. The following proposition contains a useful 'duality' principle.

PROPOSITION 6. *H_1 acts ergodically on $G\backslash H_2$ if and only if H_2 acts ergodically on $H_1\backslash G$.*

Proof. By Section 1, H_1 is not ergodic on G/H_2 if and only if there exists a Borel function f' on G/H_2 such that $f'(h_1 \cdot m) = f'(m)$ such that f' is not almost everywhere equal to a constant. Then if $f(g) = f'((gH_2))$, f is a Borel function on G and is not almost everywhere (with respect to Haar measure) a constant (see [18], Lemma 1.3). Moreover $f(h_1 \cdot gh_2) = f(g)$ for all $h_1 \in H_1$, $h_2 \in H_2$. Now by starting with $H_1\backslash G$ instead of G/H_2, we see that the existence of such an f is also the necessary and sufficient for H_2 not to act ergodically on $H_1\backslash G$.

Remark. We note that each of the conditions of the Proposition is also equivalent to the assertion that $H_1 \times H_2$ acts ergodically on G by $(h_1, h_2) \cdot g = h_1 g h_2^{-1}$ with Haar measure on G understood. (It doesn't matter whether we use left or right Haar measure since they have the same null sets.)

The rather innocent observation in Proposition 6 can be used in conjunction with the results announced in Section 3 to yield some less innocent looking facts. To be precise, let $G = SL(n, R)$ $(n \geqq 2)$ be the group of $n \times n$ unimodular real matrices, and let H_1 be the closed subgroup consisting of all elements of G such that the first column of the matrix is $(1, 0, \cdots, 0)$, but otherwise arbitrary. Let H_2 be any subgroup Γ such that $SL(n, R)/\Gamma$ has a finite G-invariant measure. Clearly H_1 is totally non-compact since G is simple and H_1 contains a non-trivial one-parameter subgroup $\exp(tx)$ with $ad(x)$ nilpotent. Then Theorem 3 (or Theorem 4) implies that H_1 is ergodic on G/Γ, and hence by Proposition 6, Γ is ergodic on $H_1\backslash G$. But we claim that $H_1\backslash G$ is isomorphic to n-space minus the origin with G acting linearly. Indeed H_1 is the stability group of the vector $(1, 0, \cdots, 0)$ in R^n under the linear action of G. Thus $H_1\backslash G$ is the orbit of this point, namely everything except the origin. Also $H_1\backslash G$ has an invariant (infinite) measure, namely Lebesgue measure. But now if Γ acts ergodically on $R^n - \{0\}$, it acts ergodically on R^n (Lebesgue measure) since $\{0\}$ is a null set.

THEOREM 6. *If Γ is a subgroup of $SL(n, R)$ such that $SL(n, R)/\Gamma$ has finite G-invariant volume, then Γ acts ergodically on R^n (ordinary linear action and Lebesgue measure understood).*

COROLLARY. *The group of $n \times n$ integral unimodular matrices, $SL(n,Z)$, acts ergodically on R^n.*

Proof. It is classic that $\Gamma = SL(n,Z)$ satisfies the hypothesis of Theorem 6 (see [6]).

Setting $n = 2$, we obtain the result announced in the introduction. One could of course give many other applications, but we content ourselves with just one which follows directly from Theorem 4. Note that the flow induced by Γ on $H\backslash G$ by right translation is equivalent topologically and measure theoretically to the left translation flow of Γ on G/H.

THEOREM 7. *Let G be a simple (connected) non-compact Lie group with finite center, and let Γ be a subgroup such that G/Γ has finite invariant volume with $\Gamma \neq G$. Then for any other closed subgroup H, Γ acts ergodically on G/H if and only if H is non-compact.*

Note that with our conditions, Γ is necessary discrete (see the proof of Theorem 4). Theorem 7 is then closely related to a classic result of C. L. Siegel ([28], Theorem 1) which says that if G is locally compact (and separable) and Γ is a discrete subgroup such that G/Γ has finite invariant volume, then for any closed subgroup H of G, Γ acts discontinuously on G/H if and only if H is compact. It is well known that ergodicity of Γ on G/H implies that 'almost all' Γ orbits are dense. In fact, it is easy to prove that the set of m in G/H such that $\Gamma \cdot m$ is not dense is of the first category and of measure zero in G/H whenever Γ is ergodic on G/H.

INSTITUTE FOR ADVANCED STUDY AND
UNIVERSITY OF CALIFORNIA, BERKELEY.

REFERENCES.

[1] L. Auslander et al., *Flows on Homogeneous Spaces*, Annals of Mathematics Studies No. 53, Princeton University Press, Princeton, 1963.

[2] ——— and L. Green, "G-induced flows on solvmanifolds" (forthcoming).

[3] A. Borel, "Groupes linéaires algébriques," *Annals of Mathematics* (2), vol. 64 (1956), pp. 20-80.

[4] ———, "Density properties for certain subgroups of semi-simple groups without compact components," *Annals of Mathematics* (2), vol. 72 (1960), pp. 179-188.

[5] ———, "Compact Clifford-Klein forms of symmetric spaces," *Topology*, vol. 2 (1963), pp. 111-122.

[6] ——— and Harish-Chandra, "Arithmetic subgroups of algebraic groups," *Annals of Mathematics* (2), vol. 75 (1962), pp. 485-535.

[7] I. Gelfand and S. Fomin, "Geodesic flows on manifolds of constant negative curvature," *Uspekhi Matematicheskikh Nauk*, vol. 7 (1952), pp. 118-137.

[8] P. Halmos, *Lectures on Ergodic Theory*, Publications of the Mathematical Society of Japan, no. 3, Mathematical Society of Japan (1956).

[9] G. Hedlund, "Fuchsian groups and transitive horocycles," *Duke Mathematical Journal*, vol. 2 (1936), pp. 530-542.

[10] ———, "Dynamics of geodesic flow," *Bulletin of the American Mathematical Society*, vol. 45 (1939), pp. 241-260.

[11] S. Helgason, *Differential Geometry and Symmetric Spaces*, Acadamic Press, New York, 1962.

[12] E. Hopf, "Statisk der Geodätische Linien in Manifaltikeiten Negativer Krümmung," *Berichte Verhandlunge der Sächsischen Akademie von Wissenschaften zu Leipzig, Mathematik-Naturforschungs Klasse*, vol. 51 (1939), pp. 261-304.

[13] N. Jacobson, "Completely reducible Lie algebras of linear transformations," *Proceedings of the American Mathematical Society*, vol. 2 (1951), pp. 105-113.

[14] ———, *Lie Algebras*, Interscience, New York, 1962.

[15] B. Kostant, "The principal three dimensional subgroup and the Betti numbers of a complex semi-simple Lie group," *American Journal of Mathematics*, vol. 81 (1959), pp. 973-1032.

[16] G. W. Mackey, *Group Representations*, Notes, University of Chicago, 1955.

[17] ———, "A theorem of Stone and von Neumann," *Duke Mathematical Journal*, vol. 16 (1949), pp. 313-326.

[18] ———, "Induced representations of locally compact groups I," *Annals of Mathematics* (2), vol. 55 (1952), pp. 101-139.

[19] ———, "Point realizations of transformation groups," *Illinois Journal of Mathematics*, vol. 6 (1952), pp. 327-335.

[20] ———, "Borel structures in groups and their duals," *Transactions of the American Mathematical Society*, vol. 85 (1957), pp. 134-165.

[21] F. I. Mautner, "Geodesic flows on symmetric Riemann spaces," *Annals of Mathematics* (2), vol. 65 (1957), pp. 416-431.

[22] V. Morosov, "On nilpotent elements of a semi-simple Lie algebra," *Akademiia Nauk SSSR* (Doklady), vol. 36 (1942), pp. 83-86.

[23] C. C. Moore, "Compactifications of symmetric spaces, II, the Cartan domains," *American Journal of Mathematics*, vol. 86 (1964), pp. 358-378.

[24] O. Parasyuk, "Horocycle flows on surfaces of constant negative curvature," *Uspekhi Matematicheskikh Nauk*, vol. 8 (1953), pp. 125-126.

[25] Seminaire Bourbaki Expose No. 206, Paris.

[26] A. Selberg, "On discontinuous groups in higher dimensional symmetric spaces," *Contributions to function theory, Tata Institute of Fundamental Research*, 1960.

[27] T. Sherman, "A weight theory for unitary representations (forthcoming).

[28] C. L. Siegel, "Discontinuous groups," *Annals of Mathematics* (2), vol. 44 (1943), pp. 674-689.

ON MAXIMAL RIEMANN SURFACES.

By H. Renggli.

1. Introduction. 1. We shall deal with imbeddings of Riemann surfaces; in particular we want to imbed a Riemann surface S into another Riemann surface T and study relations between S and T. We mean by S is *imbedded* in T, that there is a conformal mapping of S into T. A conformal mapping is understood to be a one-to-one correspondence given by an analytic function. Briefly we say that S is imbedded in T or T is an *extension* of S, and we denote it simply by $S \subset T$, identifying S and its image in T.

Very often one is interested in making the extension T maximal. One means by T is *maximal* or T is a *maximal extension* of S, that for every Riemann surface T^* the condition $(S \subset)\, T \subset T^*$ implies $T = T^*$, i.e. every conformal mapping of T into T^* maps T onto T^*. Trivially a compact Riemann surface is maximal.

2. The above mentioned concepts have been introduced by T. Rado who also gave the first example of a maximal Riemann surface which is not compact [7]. Subsequently S. Bochner proved [2], that to every Riemann surface there is a maximal extension.

3. In this paper we will prove two theorems. The first discusses when a Riemann surface is maximal and the second gives necessary and sufficient conditions in order that any maximal extension T of a Riemann surface S be compact. It will also follow that T is then topologically unique. Furthermore we shall show making use of a theorem by K. Oikawa [5] that such a compact extension T is in general not conformally unique.

4. Necessary and sufficient conditions for maximality have already been given by R. de Possel ([6], p. 15). His result can be summarized as follows. A Riemann surface T is maximal if and only if T contains no planar boundary component and every simply connected subdomain of T satisfies a certain maximality condition relative to its boundary which lies in T. But his second condition has no immediate geometric interpretation,

---

Received April 13, 1964.
Revised received August 31, 1965.

although it could be transformed using the notion of N_D-sets as has been remarked by L. Ahlfors and A. Beurling ([1], p. 124).

Instead we have preferred to give conditions which have more geometric meaning. They are based on the notions of a boundary continuum (Definition 2) and of a point-like boundary component (2.4) of a Riemann surface. The concept of a boundary continuum is also very useful in studying general questions of imbeddings (Lemmata 1 and 2).

Secondly we have to mention the work by M. Jurchescu [3], who emphasized the importance of the class of maximal Riemann surfaces for the classification problem. He also characterized maximal Riemann surfaces by properties of their boundary. However, our approach is more elementary and seems to be simpler than the one used by M. Jurchescu ([3], § 5).

2. Definitions. 1. For our study we need some classification of boundary sets of a Riemann surface S. We recall first that a subdomain R of S is planar (schlichtartig) if and only if there exists a homeomorphism of R into the Euclidean plane E. Furthermore we adopt the notions of the abstract boundary of S and of a boundary component γ of S as already given in ([4]; V, § 1). In the corresponding compactification of S every boundary component γ of S is represented by exactly one point of the abstract boundary.

We mention also that a set C separates a set A from a set B if and only if the sets A, B, C are not empty and mutually disjoint and every curve joining A with B will intersect C. Finally if G is a domain in E, ∂G will denote its boundary.

2. Using planar subdomains we shall introduce the concept of a boundary continuum of a Riemann surface.

Definition 1. *Let S be a Riemann surface. By a pair $\langle R, f\rangle$ we mean a set R and a function f where*

(i) *R is a planar proper subdomain of S such that the boundary of R which lies in S is either a closed Jordan curve C or an open Jordan arc C in S, and*

(ii) *f is a conformal mapping of R into the unit disc $D = \{z$ complex: $|z| < 1\}$, such that in the case of the Jordan curve, C corresponds to the unit circle, and in the case when C is a Jordan arc, to some subarc of it.*

Definition 2. *A Riemann surface S is said to contain a boundary continuum if and only if there exists a pair $\langle R, f\rangle$ such that the set $\partial[f(R)] - f(C)$*

contains a continuum, i. e. a closed connected set consisting of more than one point.

The set $\partial[f(R)] - f(C)$ is called a *realization* of that part of the boundary of R which does not lie in S. It could also be considered as a realization of the corresponding part of the boundary of S.

3. There are components of the boundary of S which have realizations.

Definition 3. *A union* Γ *of boundary components* γ *of a Riemann surface* S *is said to be planar (schlichtartig) if and only if there is a pair* $\langle R, f\rangle$ *such that* R *separates each* γ, $\gamma \in \Gamma$, *from* C.

In other words R is multiply connected and each γ is a boundary component of R. The image $f(\gamma)$ of each γ can be defined in an obvious way. It relates the point γ with a certain subset of D which is separated from the unit circle by $f(R)$. Or each $f(\gamma)$ is a boundary component of the multiply connected domain $f(R)$.

The image $f(\gamma)$ of each γ will also be called a realization of γ and $f(\Gamma)$ a realization of Γ. In the sequel when we speak of a realization of a planar union of boundary components we shall always mean a realization in that sense.

4. There are special planar boundary components which in every realization are represented by points.

Definition 4. *A planar union* Γ *of boundary components* γ *of a Riemann surface* S *is said to be completely point-like if and only if for any realization of* Γ *the set* $f(\Gamma)$ *is totally disconnected, i. e. every* $f(\gamma)$, $\gamma \in \Gamma$, *consists of a point.*

We shall say that a boundary component is *point-like* if and only if it is planar and completely point-like.

We recall that in the terminology of L. Ahlfors and A. Beurling [1] a compact totally disconnected set N in the Euclidean plane E is a set N_{SB} if and only if for every conformal mapping g of the complement $E - N$ into E the image $g(E - N)$ is unbounded. Let Γ be a planar union of boundary components and $\langle R, f\rangle$ a pair for Γ. Furthermore we assume that $f(\Gamma) \cup f(R)$ is a domain. Now it follows (cf. [8], p. 66) that every γ, $\gamma \in \Gamma$, is point-like if and only if every compact subset of $f(\Gamma)$ is a set N_{SB}.

3. Theorems. In this section we shall only state our two theorems. Their proofs will be given later.

THEOREM 1. *A Riemann surface T is maximal if and only if T contains no boundary continuum and T has no point-like boundary component.*

THEOREM 2. *Every maximal extension of a Riemann surface S is compact if and only if every boundary component of S is point-like. Furthermore, if the latter condition holds, any two maximal extensions have the same genus, but for genus different from zero, two such maximal extensions are not necessarily conformally equivalent.*

Therefore S has only compact maximal extensions, if and only if S is some compact Riemann surface T from which one has removed a closed totally disconnected point-set N such that the boundary N of $S=T-N$ consists of point-like boundary components. Loosely speaking we simply have removed sets N_{SB}.

It also follows that the Riemann sphere R is the unique maximal extension of a Riemann surface S if and only if S can be imbedded in R such that $R-S$ is a set N_{SB}.

4. Lemmata.

LEMMA 1. *A Riemann surface S can be imbedded in some Riemann surface T such that the interior of $T-S$ is not empty, if and only if S contains a boundary continuum.*

Proof. a) Let S contain a boundary continuum. According to Definition 2 there is a pair $\langle R, f\rangle$ such that $\partial[f(R)]-f(C)$ contains a continuum. We can now assume that $D-f(R)$ has interior points.

Because otherwise we choose a continuum δ, $\delta\subset\{\partial[f(R)]-f(C)\}$, such that the complement Δ of δ with respect to the Riemann sphere is simply connected. By the Riemann mapping theorem Δ can be mapped conformally onto the unit disc D. It is easy to see that a successive appropriately chosen application of Riemann's mapping theorem will then give such a desired pair $\langle R, f\rangle$.

$T=S\cup D$, where corresponding points in R and $f(R)$ have been identified, constitutes a Riemann surface which satisfies our requirements.

b) Let $S\subset T$ and q be an interior point of $T-S$. Let $p\in S$ and $\omega\colon\{0\leqq t\leqq 1\}\rightarrow T$ be a Jordan arc with $\omega(0)=p$ and $\omega(1)=q$. Suppose $t_0=\sup\{t\colon \omega(t)\in S\}$ and $q^*=\omega(t_0)$, then $q^*\in T-S$. Choose N to be a simply connected neighborhood of q^*, such that $q\notin N$, the boundary of N

is a Jordan curve J, $J \subset T$, and $J \cap S \neq \emptyset$. Obviously such an N exists, and $\omega(t_0 < t < 1) \cap N$ contains an open arc δ, which joins q^* with J.

Let $p^* \in J \cap S$ and Q be that component of $S \cap N$ whose closure contains p^*. Denote by B the set of open arcs β such that each β lies on J, $\beta \subset S$, and β belongs to the boundary of Q. We shall construct a subdomain R of Q whose boundary in S is an open Jordan arc C. In the case when B consists of one element β, we set $R = Q$ and $C = \beta$. Otherwise we join p^* with some point of another arc β by an open Jordan arc β^* which lies in Q. β^* partitions N into three sets: two domains and β^*. The one domain containing δ we call D_1. Put $Q_1 = D_1 \cap Q$. The boundary of Q_1 which lies in S consists of two subarcs of B joined by β^* and some other open arcs β. If the latter set is not empty, we iterate the process with Q_1 in an obvious way. Either we are through after a finite number of steps and have thus constructed a domain R whose boundary in S is an open Jordan arc C, or we are able to get an appropriately chosen sequence whose limit will give us such a desired subdomain R.

Let Δ, $R \subset \Delta$, be that simply connected domain which has C, the slit δ and a subset of J as its boundary. Let f map Δ conformally onto the unit disc D; C goes into the unit circle and δ onto an arc of it. Therefore $\partial[f(R)] - f(C)$ will contain a continuum. Hence the pair $\langle R, f \rangle$ satisfies Definitions 1 and 2, i. e. S contains a boundary continuum.

LEMMA 2. *If the Riemann surface S contains a boundary continuum, then S can be imbedded in some Riemann surface T which has infinite genus.*

Proof. By Lemma 1 we can assume there is a T_1 such that $S \subset T_1$ and q is an interior point of $T_1 - S$. Let N_1, $N_1 \subset T_1 - S$, be a simply connected neighborhood of q and suppose g_1 maps N_1 conformally onto $D_1 = \{z : |z| < 2\}$. We will modify T_1 to a Riemann surface of infinite genus, i. e. we will take out a piece of T_1 and put in its place a surface of infinite genus. We remove $D^* = \{z : |z| \leqq 1\}$ from D_1 and set $T_1^* = T_1 - D^*$, i. e. we remove from T_1 all points which are mapped by g_1 onto D^*.

Let T_2 be a Riemann surface of infinite genus and choose a simply connected subdomain N_2 of T_2. Suppose g_2 maps N_2 conformally onto the complement of D^* with respect to the Riemann sphere. Similarly we let T_2^* be the Riemann surface T_2 from which we have removed all points which are mapped onto the complement of D_1.

We put $T = T_1^* \cup T_2^*$, where we agree to identify the points of N_1 and N_2 which are mapped by g_1 and g_2 onto the same z, $z \in \{z : 1 < |z| < 2\}$. T is obviously a Riemann surface of infinite genus and $S \subset T$.

5. Proof of Theorem 1. We prove both parts of Theorem 1 by contraposition.

a) Assume T is not maximal. Then there is a T^* such that $T \subset T^*$ and $T^* - T \neq \emptyset$. Either $T^* - T$ has interior points or not. In the first case we apply Lemma 1 and hence T contains a boundary continuum. In the second case, let $p \in T^* - T$ and let N, $N \subset T^*$, be a simply connected neighborhood of p. Either the component of $N - T$ which contains p contains a continuum or consists of a point. In the former case, we choose a subcontinuum δ, $\delta \subset N - T$, such that $N - \delta$ is doubly connected, and we map N by g conformally onto D. Let h map the complement of $g(\delta)$ with respect to the Riemann sphere into the exterior of $D^* = \{z : |z| \leqq 1\}$. Now, $(T^* - \delta) \cup D^*$ can be made into a Riemann surface which has exterior points with respect to T. Hence T contains a boundary continuum by Lemma 1.

Therefore we can assume that $N - T$ is totally disconnected and hence the boundary component $\{p\}$ of T is planar. Then either $\{p\}$ is point-like or not. In the latter case, there is a pair $\langle R, f\rangle$ for $\{p\}$ such that $f(p)$ is a continuum. So we get once more by the same method as above a boundary continuum for T. In the former case, T has a point-like boundary component.

b) Let T either contain a boundary continuum or have a point-like boundary component. In the first case T is not maximal by Lemma 1. In the second case, choose a pair $\langle R, f\rangle$ for the point-like boundary component. Put $T^* = T \cup D$, where corresponding points in R and $f(R)$ have been identified. T^* is a Riemann surface such that $T \subset T^*$ and $T^* - T \neq \emptyset$.

6. Proof of Theorem 2. 1. The first part of the first statement is proved by contraposition, the second part directly.

a) Let T be maximal, not compact and $S \subset T$. If $T - S$ is not totally disconnected, we proceed as in the first part of the proof of Theorem 1. Hence S contains a boundary continuum.

Therefore we may assume that $T - S$ is totally disconnected. Since T is maximal and not compact, by use of Theorem 1, T has at least one boundary component which is not planar. But then also S must have such a boundary component.

b) By hypothesis any maximal extension T of S is compact. In the case where there is a T such that $T - S$ is not totally disconnected, we show as above that S contains a boundary continuum. This implies by Lemma 2 that S has an extension of infinite genus and we have arrived at a contradiction.

Hence $T-S$ is totally disconnected. Similarly a contradiction is obtained if not every component of $T-S$ is point-like with respect to S. But since T is compact, S has no further boundary components.

2. Let S have only point-like boundary components and let T be a maximal extension of S. We know already that T is then compact and $T-S$ totally disconnected. Hence it is almost trivial that S and T have same genus. We nevertheless outline a proof. Let Σ be a system of retrosections on T such that $T-\Sigma$ is a planar subdomain of T. Since $T-S$ is totally disconnected Σ can be chosen to belong to S. But then $S-\Sigma$ will also be a planar subdomain of S. So S and T have same genus.

In the case of genus zero, any maximal extension of S is conformally equivalent to the Riemann sphere. Hence the maximal extension is also conformally unique.

Let T be a compact Riemann surface different from the Riemann sphere. Take a simply connected subdomain G of T and a conformal mapping g of G onto the unit disc D, and choose in D a compact set which is a set N_{SB} but not a set N_D. That such sets exist has been shown by L. Ahlfors and A. Beurling ([1], Theorem 16). Let S be the Riemann surface T from which one has removed all points which are mapped by g onto N_{SB}. According to 2.4 every boundary component of S is point-like. Now it follows from a result of H. L. Royden ([9], Theorem 2), that such a Riemann surface S is not of class O_{AD}. But K. Oikawa has proved ([5], Theorem 1) that if all compact extensions of a finite genus different from zero are conformally equivalent, then the extended Riemann surface is of class O_{AD}. Hence two maximal extensions of such a Riemann surface S are not necessarily conformally equivalent.

UNIVERSITY OF NEW MEXICO.

REFERENCES.

[1] L. Ahlfors and A. Beurling, "Conformal invariants and function-theoretic null-sets," *Acta Mathematica*, vol. 83 (1950), pp. 101-129.

[2] S. Bochner, "Fortsetzung Riemannscher Flächen," *Mathematische Annalen*, vol. 98 (1928), pp. 406-421.

[3] M. Jurchescu, "Bordered Riemann surfaces," *Mathematische Annalen*, vol. 143 (1961), pp. 264-292.

[4] B. v. Kerekjarto, *Vorlesungen über Topologie I*, Grundlehren de mathematischen Wissenschaften, Springer, Berlin, 1923.

[5] K. Oikawa, "On the uniqueness of the prolongation of an open Riemann surface of finite genus," *Proceedings of the American Mathematical Society*, vol. 11 (1960), pp. 785-787.

[6] R. de Possel, "Quelques problèmes de représentation conforme," *Journal de l'école polytechnique*, II 30 (1932), pp. 1-98.

[7] T. Rado, "Ueber eine nicht fortsetzbare Riemannsche Mannigfaltigkeit," *Mathematische Zeitschrift*, vol. 20 (1924), pp. 1-6.

[8] H. Renggli, "Quasiconformal mappings and extremal lengths," *American Journal of Mathematics*, vol. 86 (1964), pp. 63-69.

[9] H. L. Royden, "On a class of null-bounded Riemann surfaces," *Commentarii Mathematici Helvetici*, vol. 34 (1960), pp. 52-66.

THE GENERALIZED SMITH CONJECTURE.

By Charles H. Giffen.*

0. Introduction and definitions. A fundamental problem in the topology and knot theory of the 3-sphere is the Smith Conjecture that no knotted simple closed curve can be the fixed point set of a non-trivial periodic transformation of the 3-sphere [5]. The Smith Conjecture may be generalized to higher dimensions by asking whether a nice knotted $(n-2)$-sphere (one which does not bound a tame disk) can be the fixed point set of a nice non-trivial periodic transformation of the n-sphere. This paper shows that, even with the usual smoothness restriction, the generalized Smith Conjecture is false in dimensions greater than three. This result requires the generalized Poincaré Conjecture in dimensions greater than four. Employing special techniques, we settle with the Smith Conjecture for dimension four in the negative for semilinear odd periodic transformations. If the generalized Poincaré Conjecture in dimension four is true, then the generalized Smith Conjecture in dimension four is also false for even periodic transformations. We do not know whether this is the best possible result in dimension four.

There is another generalization of the Smith Conjecture which says that no knotted $(n-2)$-sphere is the fixed point set of an effective action of the circle group on the n-sphere. W-Y. Hsiang [3] has proved this conjecture for differentiable actions of the circle group on the n-sphere, $n \neq 4$. He has also shown that, for $n = 4$, the complement of the fixed point set must have the homotopy type of a circle.

Our results depend heavily on the fact that the knots considered have complements that are fibered over a circle with finite bundle group in a very nice way. More explicitly, by a *relative fiber bundle*

$$\begin{tikzcd}(A,B) \arrow[r, "\mu_x"'] & (P,Q) \arrow[r, "p"'] & X\end{tikzcd}$$

we mean a fiber bundle

$$\begin{tikzcd}(A-B) \arrow[r, "\mu_x"'] & (P-Q) \arrow[r, "p"'] & X\end{tikzcd}$$

Received December 9, 1964.

* The author gratefully acknowledges the support of the National Science Foundation during the preparation of this paper.

in the sense of N. E. Steenrod [7] with the additional requirement that the natural (i.e., coordinate function induced) homeomorphism

$$\mu_x\colon (A - B) \to p^{-1}(x)$$

extend to a pair homeomorphism

$$\mu_x\colon (A, B) \to (\mathrm{Cl}\, p^{-1}(x), Q).$$

We shall always regard the group of a (relative) fiber bundle as a subgroup of the homeomorphism group of the fiber generated by the coordinate translations. Let R^n, E^n, D^n, and S^n denote real n-dimensional space, the open unit ball in R^n, the closed unit ball in R^n, and the unit sphere in R^{n+1}, respectively. With S^n regarded as (S^{k-1} join S^{n-k}), there is a canonical relative fiber bundle

$$\begin{tikzcd}(D^k, S^{k-1}) \arrow[r, "\mu_x"'] & (S^n, S^{k-1}) \arrow[r, "p"'] & S^{n-k}\end{tikzcd}$$

where the natural homeomorphism

$$\mu_x\colon (D^k, S^{k-1}) \to (S^{k-1} \text{ join } x, S^{k-1}) = (\mathrm{Cl}\, p^{-1}(x), S^{k-1})$$

is obvious. The complementary fiber bundle

$$\begin{tikzcd}E^k \arrow[r, "\mu_x"'] & (S^n - S^{k-1}) \arrow[r, "p"'] & S^{n-k}\end{tikzcd}$$

is just equivalent to the product bundle

$$E^k \to (E^k \times S^{n-k}) \to S^{n-k}.$$

The group of either bundle may be taken as the trivial group.

By identifying R^k with $R^k \times 0$ in $R^n = R^k \times R^{n-k}$, we obtain the "standard" inclusions, for example, of S^{k-1} in S^{n-1}. In this way, expressions such as (D^n, S^k) and (R^n, E^k) are unambiguous and need no further explanation. By an *n-manifold pair* (M, N), we mean that M is an n-dimensional manifold and that N is an $(n-2)$-dimensional manifold. By an *n-dimensional knot* or *n-sphere pair* (S, K), we mean that (S, K) is an n-manifold pair and that S and K are spheres. Other pairs are not necessarily of codimension two. A pair (P, Q) is *fibered* if it is the total space pair of a relative fiber bundle, as given in the preceding paragraph. Fibered knots in dimension three have been characterized by J. Stallings [6] as those whose knot groups have finitely generated commutator subgroups—formerly called Neuwirth knots.

In Section 1, we define twisted spun knots and state a fundamental theorem of E. C. Zeeman which, for our purposes, describes very well the structure of such knots as fibered knots fibered over a circle with finite bundle group. In Section 2, we deduce properties of finite cyclic coverings of the complementary space of a manifold pair which is fibered over a circle with finite bundle group. The generalized Smith Conjecture in dimensions greater than three is settled in Section 3, by an analysis of the finite cyclic branched coverings of fibered knots that are fibered over a circle with finite bundle group.

1. Twisted spun knots. Let (S,K) be a smooth n-sphere pair, $n \geqq 3$, and let $x' \in K$. Then there is a smooth embedding

$$g: (R^n, R^{n-2}) \to (S,K)$$

such that $g(0) = x'$. Let

$$(S^*, K^*) = (S,K) - g(E^n, E^{n-2}),$$

a smooth n-ball pair, be the *punctured* (S,K). Note that the boundary $(n-1)$-sphere pair

$$\mathrm{Bd}(S^*,K^*) = g(S^{n-1}, S^{n-3})$$

is trivial.

Writing R^n as $R^{n-2} \times R^2$, define *latitude-longitude* coordinates (z,ϕ) of a point of S^{n-1} as follows. The latitude z of the point is the projection of that point on R^{n-2}; the longitude ϕ of the point is the angular polar coordinate of the projection of that point on R^2 or is indeterminate if the projection on R^2 is the origin. Clearly, $z \in D^{n-2}$, and ϕ is indeterminate if and only if $z \in S^{n-3}$. Let (u,θ) be polar coordinates on D^2, and let q be an integer. Set

$$(\mathfrak{S}, \mathfrak{K}) = (S^{n-1}, S^{n-3}) \times D^2 \cup_f (S^*, K^*) \times S^1 \text{ (disjoint)},$$

where the attaching homeomorphism

$$f: (S^{n-1}, S^{n-3}) \times S^1 \to \mathrm{Bd}(S^*, K^*) \times S^1$$

is given by

$$f: (z,\phi) \times (1,\theta) \to g(z, \phi + q\theta) \times (1,\theta).$$

Then $(\mathfrak{S}, \mathfrak{K})$ is a smooth $(n+1)$-sphere pair determined only by (S,K) and $|q|$, called the *q-twist spun* (or *spin of*) (S,K); see [9].

Let r be a non-zero integer, and let

$$e: (\tilde{S}_r, \tilde{K}_r) \to (S,K)$$

be the r-fold cyclic branched covering of the smooth n-sphere pair (S,K) induced by mapping the fundamental group of the complementary space

$S-K$ onto its commutator quotient group (the infinite cyclic group Z) and then factoring out by the r-th powers to give the finite cyclic group Z_r of order r; see [1]. We adopt the convention that $(\tilde{S}_r, \tilde{K}_r)$ is oriented so that e is orientation preserving or reversing for some (fixed) orientation on (S, K) according as r is positive or negative. This is of little consequence here, but will be of use later in certain constructions. Let

$$\tilde{g}\colon (R^n, R^{n-2}) \to (\tilde{S}_r, \tilde{K}_r)$$

be a smooth embedding with

$$g^{-1}e\tilde{g}\colon (R^n, R^{n-2}) \to (R^n, R^{n-2})$$

the standard r-fold cyclic branched covering of (R^n, R^{n-2}). Let

$$(\tilde{S}_r{}^*, \tilde{K}_r{}^*) = (\tilde{S}_r, \tilde{K}_r) - \tilde{g}(E^n, E^{n-2}),$$

and note that

$$e^*\colon (\tilde{S}_r{}^*, \tilde{K}_r{}^*) \to (S^*, K^*)$$

is the r-fold cyclic branched covering of the smooth pair (S^*, K^*), where $e^* = e \mid (\tilde{S}_r{}^*, \tilde{K}_r{}^*)$. Let c, $c^* = c \mid (\tilde{S}_r{}^*, \tilde{K}_r{}^*)$ be *canonical* covering translations of $(\tilde{S}_r, \tilde{K}_r)$, $(\tilde{S}_r{}^*, \tilde{K}_r{}^*)$, respectively—that is to say, $\tilde{g}^{-1}c\tilde{g}(S^{n-1}, S^{n-3})$ is a rotation about the fixed point set S^{n-3} through an angle $2\pi/r$, and similarly for c^*. The following result of Zeeman [9] describes the structure of a twisted spun knot.

THEOREM 1.1 (Zeeman). *Let* $(\mathfrak{S}, \mathfrak{K})$ *be the q-twist spin of the smooth n-sphere pair* (S, K), $q \neq 0$, $n \geqq 3$. *Then the following statements are true:*

(1.1) *There is a relative fiber bundle*

$$(\tilde{S}_r{}^*, \mathrm{Bd}\,\tilde{S}_r{}^*) \xrightarrow[\nu e]{} (\mathfrak{S}, \mathfrak{K}) \xrightarrow[p]{} S^1$$

with group of fiber homeomorphisms precisely the group of covering translations of $(\tilde{S}_r{}^*, \tilde{K}_r{}^*)$, *generated by a canonical covering translation and being cyclic of order* $|q|$.

(1.2) *There is an action* α_ψ, $\psi \in S^1$, *of the circle group on* $(\mathfrak{S}, \mathfrak{K})$ *which maps fibers to fibers. Restricted to* $\mathfrak{K}$, *as* ψ *sweeps out* S^1 *once,* $S^{n-3} \times 0 \cup_f \emptyset$ *is left pointwise fixed and*

$$\mathfrak{K} = S^{n-3} \times D^2 \cup_f K^* \times S^1$$

is rotated about $S^{n-3} \times 0 \cup_f \emptyset$ *just once; in fact,* α_ψ *rotates* $\mathfrak{K}$ *locally through an angle* ψ *about* $S^{n-3} \times 0 \cup_f \emptyset$.

(1.3) *The action* α_ψ *on a (closed) fiber* $F_\theta = \mathrm{Cl}\, p^{-1}(\theta)$ *lying over* $\theta \in S^1$ *maps* F_θ *to* $F_{\theta+q\psi}$, *and when* $\psi' = 2\pi/q$, *the canonical covering translation* c^* *lies (via a coordinate function) over* $\alpha_{\psi'}$ *restricted to a fiber* F_θ *(which is mapped to itself by* $\alpha_{\psi'}$*). I.e., if*

$$\nu_\theta\colon (\tilde{S}_r{}^*, \mathrm{Bd}\, S_r{}^*) \to (F_\theta, \mathfrak{K})$$

is a coordinate function for F_θ, *then* $\nu_\theta c^* = (\alpha_{\psi'} \mid F_\theta)\nu_\theta$.

The last statement (1.3) is a direct consequence of Lemmas 5 and 6 of [9], although Zeeman does not include this in the statement of his main theorem. The action α_ψ on a fiber F_θ is just the translation of that fiber along the curve from θ to $\theta + q\psi$ in S^1 given by $[0,1] \to S^1$ where $t \to \theta + tq\psi$; see [7].

The identification of $\alpha_{\psi'}$ with c^* relates the action of S^1 to the canonical covering translation of $(\tilde{S}_r{}^*, \tilde{K}_r{}^*)$ which generates the cyclic bundle group of order $|q|$. This is the principal fact used to analyze the structure of the finite cyclic (branched and unbranched) coverings of twisted spun knots.

2. Cyclic coverings of fibered knots.

We have shown that knots exist in dimensions greater than three which are fibered over a circle with finite cyclic bundle group. Torus knots and links are examples of this phenomenon in dimension three. Because the knot theory of points in a 2-manifold is trivial, 2-manifold pairs fibered over a circle with finite bundle group are relatively uninteresting.

Let $(\mathfrak{M}, \mathfrak{N})$ be a smooth m-manifold pair, $m \geqq 2$, and suppose:

(2.1) *There is a relative fiber bundle*

$$\begin{tikzcd}(F, \mathrm{Bd}\, F) \arrow[r, "\mu_\theta"'] & (\mathfrak{M}, \mathfrak{N}) \arrow[r, "p"'] & S^1\end{tikzcd}$$

with finite cyclic group C *of fiber homeomorphisms.*

(2.2) *There is an action* α_ψ, $\psi \in S^1$, *of the circle group on* $(\mathfrak{M}, \mathfrak{N})$ *which maps fibers to fibers.*

(2.3) *The action* α_ψ *on a (closed) fiber* $F_\theta = \mathrm{Cl}\, p^{-1}(\theta)$ *lying over* $\theta \in S^1$ *maps* F_θ *to* $F_{\theta+q\psi}$, q *some (fixed) non-zero integer. When* $\psi' = 2\pi/q$, *the fiber homeomorphism* k *of* $(F, \mathrm{Bd}\, F)$ *lying (via a coordinate function) over* $\alpha_{\psi'}$ *restricted to a fiber* F_θ *(which is mapped to itself by* $\alpha_{\psi'}$*) generates the cyclic bundle group* C *of order* $|q|$. *I.e., if*

$$\mu_\theta\colon (F, \mathrm{Bd}\, F) \to (F_\theta, \mathfrak{N})$$

is a coordinate function for F_θ, *then* $\mu_\theta k = (\alpha_{\psi'} \mid F_\theta)\mu_\theta$.

The action α_ψ is just that induced by translating a fiber along the appropriate curve; see Section 1. Let

$$(\bar{\mathfrak{F}}, \mathrm{Bd}\,\bar{\mathfrak{F}}) = (F, \mathrm{Bd}, F) \times R^1$$

and let

$$h_r \colon (\bar{\mathfrak{F}}, \mathrm{Bd}\,\bar{\mathfrak{F}}) \to (\bar{\mathfrak{F}}, \mathrm{Bd}\,\bar{\mathfrak{F}})$$

be the homeomorphism given by

$$h_r \colon f \times x \to k^{-1}(f) \times (x + 2\pi/qr)$$

where r is a (fixed) non-zero integer. Let

$$(\mathfrak{F}, \mathrm{Bd}\,\mathfrak{F}) = (\bar{\mathfrak{F}}, \mathrm{Bd}\,\bar{\mathfrak{F}})/h_r{}^r$$

and note that the canonical map

$$(\bar{\mathfrak{F}}, \mathrm{Bd}\,\bar{\mathfrak{F}}) \to (\mathfrak{F}, \mathrm{Bd}\,\mathfrak{F})$$

is an infinite cyclic covering of $(\mathfrak{F}, \mathrm{Bd}\,\mathfrak{F})$. Let

$$\hat{e}_r \colon (\mathfrak{F}, \mathrm{Bd}\,\mathfrak{F}) \to (\mathfrak{M}, \mathfrak{N})$$

be given by

$$\hat{e}_r \colon (f \times x)/h_r{}^r \to \alpha_{rx}\mu_0(f),$$

where μ_0 is a coordinate function taking $(F, \mathrm{Bd}\,F)$ to $(F_0, \mathfrak{N})$. Clearly $\hat{e}_r \mid \mathrm{Int}\,\mathfrak{F}$ is the r-fold cyclic covering of $\mathfrak{M} - \mathfrak{N}$ induced by the composition epimorphism

$$\pi_1(\mathfrak{M} - \mathfrak{N}) \xrightarrow[p^*]{} Z \to Z_r.$$

We now factor out by the inverse images of points of $\mathfrak{N}$ by setting

$$\mathfrak{X} = \{\hat{e}_r{}^{-1}(w) \mid w \in \mathfrak{N}\},$$

and setting

$$(\bar{\mathfrak{M}}_r, \bar{\mathfrak{N}}_r) = (\mathfrak{F}, \hat{e}_r{}^{-1}(\mathfrak{N}))/\mathfrak{X}.$$

Then $\hat{e}_r$ induces a map

$$e_r \colon (\bar{\mathfrak{M}}_r, \bar{\mathfrak{N}}_r) \to (\mathfrak{M}, \mathfrak{N})$$

that is the r-fold cyclic branched covering of the pair $(\mathfrak{M}, \mathfrak{N})$ induced by the above-mentioned epimorphism.

The branched covering space $(\mathfrak{M}_r, \mathfrak{N}_r)$ clearly has the structure of a relative fiber bundle

$$(F, \mathrm{Bd}\,F) \xrightarrow[\bar{\mu}_\theta]{} (\bar{\mathfrak{M}}_r, \bar{\mathfrak{N}}_r) \xrightarrow[\bar{p}]{} S^1,$$

where $\bar{p} = pe_r$. Also, there is an action $\bar{\alpha}_\psi$, $\psi \in S^1$, of circle group on $(\bar{\mathfrak{M}}_r, \bar{\mathfrak{N}}_r)$

which maps fibers to fibers and covers α_ψ just r times in the sense that $e_r\tilde{\alpha}_\psi = \alpha_{r\psi}e_r$; it is given by

$$\tilde{\alpha}_\psi\colon (f \times x)/h_r{}^r/\mathfrak{X} \to (f \times (x+y))/h_r{}^r/\mathfrak{X}$$

where $y \equiv \psi \pmod{2\pi}$. This may be summarized as follows.

PROPOSITION 2.1. *Let* $(\mathfrak{M},\mathfrak{N})$ *be a smooth m-manifold pair,* $m \geqq 2$, *satisfying* (2.1), (2.2), (2.3) *above, and let*

$$e_r\colon (\tilde{\mathfrak{M}}_r, \tilde{\mathfrak{N}}_r) \to (\mathfrak{M},\mathfrak{N})$$

be the r-fold cyclic branched covering of $(\mathfrak{M},\mathfrak{N})$, $r \neq 0$. *Then the following statements are true:*

(2.1′) *There is a branched fiber bundle*

$$(F, \operatorname{Bd} F) \xrightarrow[\tilde{\mu}_\theta]{} (\tilde{\mathfrak{M}}_r, \tilde{\mathfrak{N}}_r) \xrightarrow[\tilde{p}]{} S^1,$$

where $\tilde{p} = pe_r$, *with group* $\tilde{C} \approx C^r \approx Z_{q/(q,r)}$ *of fiber homeomorphisms.*

(2.2′) *There is an action* $\tilde{\alpha}_\psi$, $\psi \in S^1$, *of the circle group on* $(\tilde{\mathfrak{M}}_r, \tilde{\mathfrak{N}}_r)$ *which maps fibers to fibers satisfying* $e_r\tilde{\alpha}_\psi = \alpha_{r\psi}e_r$.

(2.3′) *The action* $\tilde{\alpha}_\psi$ *on a (closed) fiber* $G_\theta = \operatorname{Cl} p^{-1}(\theta)$ *lying over* $\theta \in S^1$ *maps* G_θ *to* $G_{\theta+q\psi}$, *and when* $\psi' = 2\pi/q$, *the fiber homeomorphism* k^r *lies (via a coordinate function) over* $\tilde{\alpha}_{\psi'}$ *restricted to a fiber* G_θ *(which is mapped to itself by* $\alpha_{\psi'}$*); thus* k^r *generates* $\tilde{C}$. *I. e., if*

$$\tilde{\mu}_\theta\colon (F, \operatorname{Bd} F) \to (G_\theta, \mathfrak{N}_r)$$

is a coordinate function for G_θ, *then* $\tilde{\mu}_\theta k^r = (\alpha_{\psi'} \mid G_\theta)\mu_\theta$.

Since $(\tilde{\mathfrak{M}}_r, \tilde{\mathfrak{N}}_r) \cong (\tilde{\mathfrak{M}}_{-r}, \tilde{\mathfrak{N}}_{-r})$ is obvious, and $k^q = 1$, we deduce easily the following corollary.

COROLLARY 2.2. *Under the hypotheses of Proposition* 2.1, *if r, s, and i are non-zero integers such that* $r \equiv \pm s \pmod{q}$, *then*

$$\tilde{\mathfrak{M}}_{iq} - \tilde{\mathfrak{N}}_{iq} \cong \operatorname{Int} F \times S^1, \tag{2.4}$$

$$\tilde{\mathfrak{M}}_r - \tilde{\mathfrak{N}}_r \cong \tilde{\mathfrak{M}}_s - \tilde{\mathfrak{N}}_s. \tag{2.5}$$

These relations among the complementary spaces $\tilde{\mathfrak{M}}_i - \tilde{\mathfrak{N}}_i$ may be extended to certain relations among the pairs $(\tilde{\mathfrak{M}}_i, \tilde{\mathfrak{N}}_i)$ in the case of twisted spun knots. We describe this in the next section.

3. Branched cyclic coverings of twisted spun knots. Let $(\mathfrak{S}, \mathfrak{K})$ be the q-twist spin of a smooth n-sphere pair (S, K), $n \geqq 3$, and let $(\tilde{\mathfrak{S}}_i, \tilde{\mathfrak{K}}_i)$ denote the i-fold cyclic branched covering of $(\mathfrak{S}, \mathfrak{K})$ constructed as in the previous section, $i \neq 0$. Let r, s be non-zero integers such that $r \equiv \pm s$ $(\operatorname{mod} q)$; then by (2.5) of Corollary 2.2,

$$\tilde{\mathfrak{S}}_r - \tilde{\mathfrak{K}}_r \cong \tilde{\mathfrak{S}}_s - \tilde{\mathfrak{K}}_s.$$

Let $\tilde{\mathfrak{A}}_r$ and $\tilde{\mathfrak{A}}_s$ denote the closed complements of regular neighborhoods of $\tilde{\mathfrak{K}}_r$ and $\tilde{\mathfrak{K}}_s$ in $\tilde{\mathfrak{S}}_r$ and $\tilde{\mathfrak{S}}_s$ respectively; then $\tilde{\mathfrak{A}}_r \cong \tilde{\mathfrak{A}}_s$, and this homeomorphism induces isomorphisms

$$\pi_1(\tilde{\mathfrak{A}}_r) \approx \pi_1(\tilde{\mathfrak{A}}_s)$$
$$\langle i_r \pi_1(\operatorname{Bd} \tilde{\mathfrak{A}}_r) \rangle \approx \langle i_s \pi_1(\operatorname{Bd} \tilde{\mathfrak{A}}_s) \rangle$$

where i_r, i_s are the inclusion induced homomorphisms (monomorphisms in this case), and $\langle \cdot \cdot \cdot \rangle$ denotes the smallest normal subgroup containing $\cdot \cdot \cdot$. Since $n \geqq 3$, regular neighborhoods of $\tilde{\mathfrak{K}}_r$ and $\tilde{\mathfrak{K}}_s$ are simply connected. By the van Kampen theorem,

$$\pi_1(\tilde{\mathfrak{S}}_r) \approx \pi_1(\tilde{\mathfrak{A}}_r) / \langle i_r \pi_1(\operatorname{Bd} \tilde{\mathfrak{A}}_r) \rangle$$
$$\pi_1(\tilde{\mathfrak{S}}_s) \approx \pi_1(\tilde{\mathfrak{A}}_s) / \langle i_s \pi_1(\operatorname{Bd} \tilde{\mathfrak{A}}_s) \rangle,$$

and by the above remarks, we conclude that $\pi_1(\tilde{\mathfrak{S}}_r) \approx \pi_1(\tilde{\mathfrak{S}}_s)$.

To calculate the homology groups of $\tilde{\mathfrak{S}}_r$, consider the following part of the homology sequence of the pair $(\tilde{\mathfrak{S}}_r, \tilde{\mathfrak{A}}_r)$:

$$H_{i+1}(\tilde{\mathfrak{S}}_r, \tilde{\mathfrak{A}}_r) \to H_i(\tilde{\mathfrak{A}}_r) \to H_i(\tilde{\mathfrak{S}}_r) \xrightarrow[\gamma_i]{} H_i(\tilde{\mathfrak{S}}_r, \tilde{\mathfrak{A}}_r) \xrightarrow[\delta_i]{} H_{i-1}(\tilde{\mathfrak{A}}_r).$$

By excision, $H_i(\tilde{\mathfrak{S}}_r, \tilde{\mathfrak{A}}_r) \approx H_i(D^2 \times S^{n-1}, S^1 \times S^{n-1})$ and is 0 in all dimensions except 2 and $n+1$ where it is Z. The homomorphism γ_i is therefore the zero homomorphism in all dimensions except possibly 2 and $n+1$; it is easy to see the γ_{n+1} is an isomorphism and that δ_{n+1} is the zero homomorphism. A generator of $H_2(\tilde{\mathfrak{S}}_r, \tilde{\mathfrak{A}}_r)$ is mapped by δ_2 to a generator of a free summand of $H_1(\tilde{\mathfrak{A}}_r)$, and therefore γ_2 is the zero homomorphism. Clearly, $H_1(\tilde{\mathfrak{S}}_r) \oplus Z \approx H_1(\tilde{\mathfrak{A}}_r)$, and from the above facts we see that $H_{n+1}(\tilde{\mathfrak{S}}_r) \approx Z$ and, for all other i, $H_i(\tilde{\mathfrak{S}}_r) \approx H_i(\tilde{\mathfrak{A}}_r)$. We now see that the homology groups of $\tilde{\mathfrak{S}}_r$ depend only upon the homology groups of $\tilde{\mathfrak{A}}_r$. Therefore, $H_i(\tilde{\mathfrak{S}}_r) \approx H_i(\tilde{\mathfrak{S}}_s)$ for all i. Summarizing the last two paragraphs, we have the following theorem.

THEOREM 3.1. *Let r, s be non-zero integers such that $r \equiv \pm s$ $(\operatorname{mod} q)$. Then $\tilde{\mathfrak{S}}_r$ and $\tilde{\mathfrak{S}}_s$ have isomorphic fundamental groups and isomorphic homology groups. The homology groups of $\tilde{\mathfrak{S}}_r$ and $\tilde{\mathfrak{S}}_r - \tilde{\mathfrak{K}}_r$ are isomorphic in all dimensions except 1 and $n+1$; $H_1(\tilde{\mathfrak{S}}_r) \oplus Z \approx H_1(\tilde{\mathfrak{S}}_r - \tilde{\mathfrak{K}}_r)$ and $H_{n+1}(\tilde{\mathfrak{S}}_r) \approx Z$.*

If $r \equiv \pm 1$ $(\operatorname{mod} q)$, choose a smooth $(n+1)$-cell in $\tilde{\mathfrak{S}}_r$ and collapse the complement in $\tilde{\mathfrak{S}}_r$ of its interior to a point; map the resulting $(n+1)$-sphere homeomorphically onto S^{n+1}, and call the composition of these operations the map f. The induced homomorphism of homology in dimension $n+1$ is clearly an isomorphism, and by Theorem 3.1, since all other homology groups in dimensions greater than one are zero, f induces isomorphisms of homology in these dimensions also. Moreover, $\tilde{\mathfrak{S}}_r$ is simply connected, and by the corollary to Theorem 3 of [8], f is therefore an homotopy equivalence. Applying the generalized Poincaré Conjecture, we have the following.

COROLLARY 3.2. *If* $(\mathfrak{S},\mathfrak{K})$ *is the q-twist spin,* $q \neq 0$, *of the smooth n-sphere pair* (S,K), $n \geqq 3$, *and if* $(\tilde{\mathfrak{S}}_r,\tilde{\mathfrak{K}}_r)$, *where r is a non-zero integer such that* $r \equiv \pm 1 \pmod{q}$, *is the r-fold cyclic branched covering of* $(\mathfrak{S},\mathfrak{K})$, *then* $\tilde{\mathfrak{S}}_r$ *is an* $(n+1)$*-sphere when* $n \geqq 4$ *and is an homotopy sphere when* $n=3$.

We can now prove that the generalized Smith Conjecture, even in the smooth case, is false for transformations of all possible periods in dimensions greater than four, and, modulo Poincaré Conjecture, that it is false for all possible periods in dimension four. For a given period $r>1$, it is easy to find an integer $q>1$ such that the hypothesis of Corollary 3.2 is satisfied ($r+1$ will work) and such that $\mathfrak{S}-\mathfrak{K}$ is topologically different from $S^{n+1}-S^{n-1}$. Then by Corollary 2.2, $\tilde{\mathfrak{S}}_r-\tilde{\mathfrak{K}}_r$ is also topologically different from $S^{n+1}-S^{n-1}$, and therefore $(\tilde{\mathfrak{S}}_r,\tilde{\mathfrak{K}}_r)$ is a non-trivial (knotted) smooth $(n+1)$-sphere pair. Let τ be a canonical covering translation of $(\tilde{\mathfrak{S}}_r,\tilde{\mathfrak{K}}_r)$; then τ generates the group of covering translations of $(\tilde{\mathfrak{S}}_r,\tilde{\mathfrak{K}}_r)$, is periodic of order r, and leaves $\tilde{\mathfrak{K}}_r$ pointwise fixed. This gives the desired result.

THEOREM 3.3. *Let* $m>4$ *and* $r>1$ *be integers. Then there exists a smooth m-sphere pair* $(\mathcal{S},\mathcal{K}) \not\cong (S^m,S^{m-2})$ *and a smooth transformation of* $(\mathcal{S},\mathcal{K})$ *of period r which leaves* $\mathcal{K}$ *pointwise fixed. If the 4-dimensional Poincaré Conjecture is true, the same holds true for* $m=4$.

For the present, we only know that $\tilde{\mathfrak{S}}_r$, in the above construction, is an homotopy 4-sphere when $n+1=4$. Using techniques due to Gluck [2], we can salvage Theorem 3.3 for $m=4$ and all odd periods r *without* the 4-dimensional Poincaré Conjecture. First, we need an alternate description of $(\tilde{\mathfrak{S}}_r,\tilde{\mathfrak{K}}_r)$. To do this, we adopt the notation of Section 2, with $(\mathfrak{S},\mathfrak{K})$ replacing $(\mathfrak{M},\mathfrak{N})$, which is allowed by Theorem 1.1; here, $m=n+1 \geqq 4$.

Let r, s be non-zero integers such that $r \equiv s \pmod{q}$. Since $(\tilde{\mathfrak{S}}_r,\tilde{\mathfrak{K}}_r) \cong (\tilde{\mathfrak{S}}_{-r},\tilde{\mathfrak{K}}_{-r})$ is obvious, in view of the remarks made in Section 1, $r \equiv \pm s \pmod{q}$ is equivalent to the above assumption. Note that $k^q=1$; then

$$\begin{aligned} h_r{}^r(f\times x) &= k^{-r}(f)\times(x+2\pi r/qr) \\ &= k^{-s}(f)\times(x+2\pi s/qs) \\ &= h_s{}^s(f\times x), \end{aligned}$$

and therefore $h_r{}^r = h_s{}^s$; whence

$$(\tilde{\mathfrak{F}}, \operatorname{Bd}\tilde{\mathfrak{F}})/h_r{}^r = (\mathfrak{F}, \operatorname{Bd}\mathfrak{F}) = (\tilde{\mathfrak{F}}, \operatorname{Bd}\tilde{\mathfrak{F}})/h_s{}^s.$$

We may now obtain $\tilde{\mathfrak{S}}_r$ or $\tilde{\mathfrak{S}}_s$ by attaching a copy of $\mathfrak{K} \times D^2$ to $\mathfrak{F}$ along their boundaries in a certain way. Let D^2 be given polar coordinates (u, θ). The attaching map ϕ_r is given by

$$\phi_r\colon \mu_0(f) \times (1, \theta) \to (\mu_0{}^{-1}\alpha_{-r\theta/q}\mu_0(f) \times \theta/q)/h_r{}^r,$$

where $f \in \operatorname{Bd} F$ and $0 \leqq \theta \leqq 2\pi$; the attaching map ϕ_s is defined similarly. Composing ϕ_r with $\hat{e}_r$, we see that

$$\hat{e}_r\phi_r\colon \mu_0(f) \times (1, \theta) \to \mu_0(f),$$

and the above-mentioned connected sum corresponds to factoring out by the circles $\hat{e}_r{}^{-1}(w)$, $w \in \mathfrak{K}$, as was done in Section 2. Therefore we conclude

$$\tilde{\mathfrak{S}}_r \cong \mathfrak{K} \times D^2 \cup_{\phi_r} \mathfrak{F}$$

and, similarly,

$$\tilde{\mathfrak{S}}_s \cong \mathfrak{K} \times D^2 \cup_{\phi_s} \mathfrak{F}.$$

Now let ϕ be the transformation of $\mathfrak{K} \times S^1$ given by

$$\phi\colon \mu_0(f) \times (1, \theta) \to \alpha_\theta\mu_0(f) \times (1, \theta),$$

and let $r = s + iq$. Then

$$\begin{aligned} \phi_r\phi^i(\mu_0(f) \times (1, \theta)) &= (\mu_0{}^{-1}\alpha_{i\theta - r\theta/q}\mu_0(f) \times \theta/q)/h_r{}^r \\ &= (\mu_0{}^{-1}\alpha_{-s\theta/q}\mu_0(f) \times \theta/q)h_s{}^s \\ &= \phi_s(\mu_0(f) \times (1, \theta)), \end{aligned}$$

and therefore $\phi_r\phi^i = \phi_s$, or $\phi^i = \phi_r{}^{-1}\phi_s$.

If $n + 1 = m = 4$ and if i is even, then by Theorem 13.2 of [2], ϕ^i is isotopic to the identity, whence ϕ_r is isotopic to ϕ_s. By Theorem 2.1 of [2],

$$(\tilde{\mathfrak{S}}_r, \tilde{\mathfrak{K}}_r) \cong (\tilde{\mathfrak{S}}_s, \tilde{\mathfrak{K}}_s),$$

and we obtain the following theorem.

THEOREM 3.4. *If $n = 3$, and if $(\mathfrak{S}, \mathfrak{K})$ is the q-twist spin, $q \neq 0$, of the smooth n-sphere pair (S, K), then, if r, s are non-zero integers such that $r \equiv \pm s \pmod{2q}$, the r-fold and s-fold cyclic branched coverings of $(\mathfrak{S}, \mathfrak{K})$ are semi-linearly equivalent. The same holds true for those values of $n > 3$ for which ϕ^2 is isotopic to the identity.*

To settle the generalized Smith Conjecture in dimension four for transformations of odd period, take $(\mathfrak{S}, \mathfrak{K})$ to be non-trivial (knotted) 2-twist spun 3-sphere pair. Then, for every odd integer r, Theorem 3.4 tells us that the r-fold cyclic branched covering of $(\mathfrak{S}, \mathfrak{K})$ is equivalent to $(\mathfrak{S}, \mathfrak{K})$. As before, a canonical covering translation is of period r and leaves $\mathfrak{K}$ pointwise fixed. We state this as follows.

THEOREM 3.5. *There exists a knotted smooth 4-sphere pair* $(\mathcal{S}, \mathcal{K})$ *which admits smooth transformations of arbitrary odd period leaving* $\mathcal{K}$ *pointwise fixed.*

Observe that, at the cost of excluding even periods, we obtain many counterexamples to the generalized Smith Conjecture in dimension four each of which works for *all* odd periods.

There is yet another way of approaching the problem of whether finite cyclic branched coverings of a twisted spun sphere pair are equivalent to the twisted spun sphere pair. Let $n \geqq 3$; since $\alpha_\psi \mid \mathfrak{K}$ rotates $\mathfrak{K}$ about an unknotted $(n-3)$-subsphere, $\mu_0^{-1}\alpha_\psi\mu_0 \mid \operatorname{Bd} F$ rotates $\operatorname{Bd} F$ about an unknotted $(n-3)$-subsphere. If $\mu_0^{-1}\alpha_{i\psi}\mu_0 \mid \operatorname{Bd} F$ extends to an action β_ψ of the circle group on F, let

$$T\colon (\mathfrak{F}, \operatorname{Bd} \mathfrak{F}) \to (\mathfrak{F}, \operatorname{Bd} \mathfrak{F})$$

be the transformation given by

$$T\colon (f \times x)/h_r{}^r \to (\beta_{qx}(f) \times x)/h_r{}^r.$$

Composing this with ϕ_r and setting $r = s + iq$, we have

$$\begin{aligned} T\phi_r(\mu_0(f) \times (1, \theta)) &= (\mu_0{}^{-1}\alpha_{i\theta - r\theta/q}\mu_0(f) \times \theta/q)/h_r{}^r \\ &= \phi_s(\mu_0(f) \times (1, \theta)), \end{aligned}$$

and therefore $T\phi_r = \phi_s$. Regarding $(\tilde{\mathfrak{S}}_r, \tilde{\mathfrak{K}}_r)$ as

$$(\mathfrak{K} \times D^2, \mathfrak{K} \times 0) \cup_{\phi_r} (\mathfrak{F}, \emptyset)$$

and $(\tilde{\mathfrak{S}}_s, \tilde{\mathfrak{K}}_s)$ as the corresponding connected sum with ϕ_s, T extends to an homeomorphism

$$T^*\colon (\tilde{\mathfrak{S}}_r, \tilde{\mathfrak{K}}_r) \to (\tilde{\mathfrak{S}}_s, \tilde{\mathfrak{K}}_s)$$

by setting $T^* = 1$ on $\mathfrak{K} \times D^2$ and $T^* = T$ on $\mathfrak{F}$.

In this way, it may be possible to obtain pair equivalences for the various finite cyclic branched coverings of $(\mathfrak{S}, \mathfrak{K})$. If $(M, \operatorname{Bd} M)$ is an n-manifold with a single $(n-1)$-sphere boundary, and if there is an action of the circle group on $(M, \operatorname{Bd} M)$ which rotates $\operatorname{Bd} M$ about an unknotted $(n-3)$-subsphere exactly i times, i a positive integer, we say that $(M, \operatorname{Bd} M)$ *can be i-spun* (in the sense of Gluck). If $i = 1$, we say that $(M, \operatorname{Bd} M)$ *can be spun.* In this language, the above result may be stated as follows.

THEOREM 3.6. *If* $(\tilde{S}_q{}^*, \mathrm{Bd}\,\tilde{S}_q{}^*)$ *can be i-spun,* $i \geqq 1$, *and if* r, s *are non-zero integers such that* $r \equiv \pm s \pmod{iq}$, *then* $(\tilde{\mathfrak{S}}_r, \tilde{\mathfrak{K}}_r) \cong (\tilde{\mathfrak{S}}_s, \tilde{\mathfrak{K}}_s)$.

If $n = 3$ and $\tilde{S}_q$ is a lens space, then according to [9] $(\tilde{S}_q{}^*, \mathrm{Bd}\,\tilde{S}_q{}^*)$ can be spun. As luck would have it, lens spaces do occur as the branched double coverings of certain knots—for example, the torus knot of type $2j+1, 2$ has the lens space of type $2j+1, 1$ as its branched double covering. This fact may be used to give an alternate proof for the falsity of the generalized Smith Conjecture in dimension four for odd periods as was done above. Unfortunately, it is not known whether $(\tilde{S}_q{}^*, \mathrm{Bd}\,\tilde{S}_q{}^*)$ can be spun when $q > 2$ for any 3-dimensional knot (S, K). If this is true for some odd q, then transformations of S^4 of even period leaving a knotted smooth 2-sphere pointwise fixed can be constructed.

PRINCETON UNIVERSITY AND
THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] R. H. Fox, "Covering spaces with singularities," *Algebraic Geometry and Topology*, Princeton Mathematics Series 12, Princeton University Press, Princeton (1957).

[2] H. R. Gluck, "The embeddings of two-spheres in the four-sphere," *Transactions of the American Mathematical Society*, vol. 104 (1962), pp. 308-333.

[3] W-Y, Hsiang, "On the unknottedness of the fixed point set of differentiable circle group actions on spheres—P. A. Smith conjecture," *Bulletin of the American Mathematical Society*, vol. 70 (1964), pp. 678-680.

[4] S. Smale, "Generalized Poincaré's conjecture in dimensions greater than four," *Annals of Mathematics*, vol. 74 (1961), pp. 391-406.

[5] P. A. Smith, "Transformations of finite period, II," *Annals of Mathematics*, vol. (2) 40 (1939), pp. 690-711.

[6] J. Stallings, "On fibering certain 3-manifolds," *Topology of 3-Manifolds and Related Topics*, Prentice Hall (1962), pp. 95-100.

[7] N. E. Steenrod, *The Topology of Fiber Bundles*, Princeton Mathematics Series 14, Princeton University Press (1951).

[8] J. H. C. Whitehead, "On the homotopy type of *ANR*'s," *Bulletin of the American Mathematical Society*, vol. 54 (1948), pp. 1133-1145.

[9] E. C. Zeeman, "Twisting spun knots," to appear in *Transactions of the American Mathematical Society*.

PROJECTIVE REPRESENTATION OF ALGEBRAIC LINEAR GROUPS OF TRANSFORMATIONS.[1]

By T. KAMBAYASHI.

Let $\mathfrak{g}$ be an algebraic linear group, not necessarily connected, operating on a complete normal variety V as a group of biregular automorphisms of V. Let V^* be a birational model of V in a projective space P^n, not contained in a hyperplane of P^n. The biregular automorphisms of P^n form the well-known algebraic linear group $PL(n)$, and those elements of $PL(n)$ which induce automorphisms of V^* form a subgroup $PL(n, V^*)$ of $PL(n)$. *The purpose of the present paper, then, is to prove that* V^* *can be so chosen that there is a canonical rational homomorphism of* $\mathfrak{g}$ *into* $PL(n, V^*)$ *whose kernel consists of the elements of* $\mathfrak{g}$ *which act as identity on* V; *furthermore, if* V *admits a biregular projective model,* V^* *can be chosen among such models.* Thus, one may say that the action of $\mathfrak{g}$ on a complete (resp. projective) normal variety can be identified with the action of a group of projective transformations if one imbeds V birationally (resp. biregularly) in an appropriate projective space.[2] For a more precise formulation of these statements, see 2.5 and 2.6 below.

Basic theorems and definitions employed in this article are found in Rosenlicht [7], Weil [9], [10]; Matsusaka's important work [5] is responsible for the main idea of the present work (to wit, represent automorphisms by their action on a divisor), although no result in [5] is used here explicitly. One should also mention the fundamental works of Barsotti (cf. [1] for instance) and Borel's paper [2]. Also, see Matsumara [4] for related results.

Throughout, an algebraically closed ground field k is fixed; all varieties, divisors, mappings and functions will be tacitly assumed rational over k, unless otherwise stated. If V is a variety and $K \supseteq k$ a field of definition for $V, K(V)$ will denote the abstract field of functions on V rational over K.

1. Construction of projective representations of certain algebraic groups. Let V be a complete normal variety, and D a divisor on V. Let

---

Received May 18, 1964.

[1] Presented to the Society at the 69th Summer Meeting, Amherst, Mass., August 1964.

[2] When $\mathfrak{g}$ is connected and semi-simple, and $V = \mathfrak{g}/B$ (B = a Borel subgroup), more refined results are obtained; cf. Exposé 15, Séminaire C. Chevalley, 1956/57.

$\mathfrak{g}$ be an algebraic group which operates regularly on V (see [7]). We have a morphism $m\colon \mathfrak{g} \times V \to V$ which defines the operation of $\mathfrak{g}$ on V; for each point x on V and for each $\sigma \in \mathfrak{g}$, we write σx for $m(\sigma, x)$. The σ-transform of D will be denoted by D^{σ}, and we shall agree that $D^{\rho\tau} = (D^{\tau})^{\rho}$ for $\rho \in \mathfrak{g}$, $\tau \in \mathfrak{g}$. We let $\sigma \in \mathfrak{g}$ operate on the functions on V by defining $f^{\sigma}(x) = f(\sigma^{-1}x)$ for every function f, where x is a generic point of V over $k(\sigma)$. It follows that $f^{\rho\tau} = (f^{\tau})^{\rho}$ for $\rho \in \mathfrak{g}$, $\tau \in \mathfrak{g}$. Denoting by $\operatorname{div}(f)$ the divisor of the function f, one can easily verify that $\operatorname{div}(f^{\sigma}) = (\operatorname{div}(f))^{\sigma}$.

Let V, D, $\mathfrak{g}$ be as above. Let $L(D)$ be the finite dimensional vector space over the constants of functions f on V with $\operatorname{div}(f) + D \geqq 0$ [10, p. 347]. For a field $K \supseteq k$, we denote by $L(D; K)$ the K-vector space of those functions in $L(D)$ which are rational over K. If f, g are functions in $L(D; K)$ (resp. in $L(D)$) such that $cf = g$ for a suitable non-zero $c \in K$ (resp. for a suitable non-zero constant c), then we write $f \equiv g \pmod{K^*}$ (resp. $f \equiv g \pmod{\Omega^*}$); this is an equivalence relation which enables us to regard the quotient structure as $P^n(K)$, *viz.* the projective n-space over K (resp. as P^n), where $n+1 = \dim L(D)$. Without fear of confusion, we usually omit the qualifications $\bmod K^*$ or $\bmod \Omega^*$.

With the same notations and assumptions, suppose that, for an element $\sigma \in \mathfrak{g}$, the divisor D^{σ} is linearly equivalent to D [10, p. 271]; notation for the situation is $D^{\sigma} \sim D$. Then, there is a function g_{σ} on V, unique up to non-zero constant factor, such that $D^{\sigma} = D + \operatorname{div}(g_{\sigma})$. We note that, if $f \in L(D)$,

$$\begin{aligned}\operatorname{div}(f^{\sigma}g_{\sigma}) + D &= \operatorname{div}(f^{\sigma}) + \operatorname{div}(g_{\sigma}) + D \\ &= (\operatorname{div}(f))^{\sigma} + D^{\sigma} = (\operatorname{div}(f) + D)^{\sigma} \geqq 0.\end{aligned}$$

Therefore, the assignment of $f^{\sigma}g_{\sigma}$ to each $f \in L(D)$ is a linear transformation of $L(D)$ which is unique up to non-zero constant multiplicative factor once σ is given. In view of this fact, we define an element $T_D(\sigma) \in PL(n)$ (where $n+1 = \dim L(D)$) by the formula: $T_D(\sigma)f \equiv f^{\sigma}g_{\sigma}$ for every $f \in L(D)$.

1.1. PROPOSITION. *Let V be a complete, normal variety on which an algebraic group $\mathfrak{g}$ operates regularly. Let D be a divisor on V, and assume that $D^{\sigma} \sim D$ for every $\sigma \in \mathfrak{g}$. Then, the mapping $\sigma \to T_D(\sigma)$ is a projective representation of $\mathfrak{g}$.*

Proof. Let $\sigma, \tau \in \mathfrak{g}$, and $f \in L(D)$. Then, $T_D(\sigma\tau)f \equiv f^{\sigma\tau}g_{\sigma\tau}$ with $D^{\sigma\tau} \sim D + \operatorname{div}(g_{\sigma\tau})$. But

$$\begin{aligned}D^{\sigma\tau} &= (\mathrm{D} + \operatorname{div}(g_{\tau}))^{\sigma} = D^{\sigma} + \operatorname{div}(g_{\tau}{}^{\sigma}) \\ &= D + \operatorname{div}(g_{\sigma}) + \operatorname{div}(g_{\tau}{}^{\sigma}) = D + \operatorname{div}(g_{\sigma}g_{\tau}{}^{\sigma}).\end{aligned}$$

Hence, $g_{\sigma\tau} \equiv g_\sigma g_\tau{}^\sigma$. On the other hand, $T_D(\tau)f \equiv f^\tau g_\tau$, and, therefore, $T_D(\sigma)T_D(\tau)f \equiv (f^\tau g_\tau)^\sigma g_\sigma = f^{\sigma\tau} g_\sigma g_\tau{}^\sigma$. Thus, $T_D(\sigma\tau) = T_D(\sigma)\cdot T_D(\tau)$ is proved.

In order to prove that the mapping is a morphism of varieties, we may assume $\mathfrak{g}$ to be connected; let γ be a generic point of $\mathfrak{g}$ over k, and set $K = k(\gamma)$. Then, γ is a K-automorphism of $K(V)$ and g_γ is rational over K. It follows then that $f^\gamma g_\gamma \in L(D;K)$ for every $f \in L(D;K)$. Therefore, $T_D(\gamma)$ may be considered as an automorphism of $P^n(K)$ where $n+1 = \dim L(D)$ and, as such, $T_D(\gamma)$ is rational over K. This proves $k(T_D(\gamma)) \subseteq K = k(\gamma)$. Q. E. D.

In accordance with classical terminology, we shall call a divisor D on V *simple* if D is positive and if $L(D)$ generates the function field of V over the constants.

1.2. PROPOSITION. *Notations and assumptions being the same as in* 1.1, *assume further that D is simple, and that $\mathfrak{g}$ acts on V effectively* (resp. *faithfully*). *Then, the homomorphism T_D is a purely inseparable isogeny* (resp. *an isomorphism*).

Proof. It is enough to prove that the kernel of T_D is $\{\epsilon\}$ ($\epsilon =$ the identity element of $\mathfrak{g}$). Suppose that $T_D(\sigma) =$ identity transformation on $P^n(k)$. This means that $f^\sigma g_\sigma \equiv f$ for every $f \in L(D;k)$. Since D is positive, $1 \in L(D;k)$ and, hence, $g_\sigma \equiv 1$ or $g_\sigma \in k$. Thus we have $D^\sigma = D$ and $f^\sigma = c(f)f$, where $c(f) \in k$ depends only on $f \in L(D;k)$. But the eigenvalue $c(f)$ must equal 1 for every $f \in L(D;k)$. Indeed, $c(a) = 1$ for every constant $a \in k$; and, if f_1, f_2 are linearly independent over k, we get

$$\begin{aligned}(f_1+f_2)^\sigma &= f_1{}^\sigma + f_2{}^\sigma = c(f_1)f_1 + c(f_2)f_2 \\ &= c(f_1+f_2)(f_1+f_2) = c(f_1+f_2)f_1 + c(f_1+f_2)f_2.\end{aligned}$$

which implies $c(f_1) = c(f_2) = c(f_1+f_2)$. Thus, by taking a basis

$$\{f_0 = 1, f_1, \cdots, f_d\}$$

of $L(D;k)$, we find that $c(f_i) = 1$ for $0 \leqq i \leqq d$. Hence, $c(f) = 1$ for every $f \in L(D;k)$. Since this vector space generates $k(V)$ over k we deduce that σ fixes $k(V)/k$ elementwise and must therefore be the identity transformation on V. Q. E. D.

2. The case when $\mathfrak{g}$ is linear; conclusions. In this section, we will first show that *connected* algebraic linear groups do satisfy the basic assumption of 1.1 and 1.2. The case when $\mathfrak{g}$ is disconnected will then be brought under the control of 1.1 and 1.2 by means of a proper choice of the divisor D (see 2.4). Finally, we will draw conclusions (2.5 and 2.6) announced at the beginning of the article.

2.1. LEMMA. *Let V, W be abstract varieties, and $f: V \to W$ a rational mapping. Let $\mathfrak{D}$ be the set of points on V at which f is regular. Then, $\mathfrak{D}$ is open; furthermore, the set $f(\mathfrak{D})$ contains an open subset of the geometric image of V by f.*

The lemma is well-known. For instance, the first assertion occurs in [10, Chap. VII, § 6, p. 195] and the second in [9, Appendix, Prop. 10, p. 384].

2.2. PROPOSITION. *Let V be a complete normal variety on which a connected algebraic linear group $\mathfrak{g}$ operates regularly. Let D be a divisor on V. Then, D^σ is linearly equivalent to D for each $\sigma \in \mathfrak{g}$.*

Proof. We draw upon the fact that connected algebraic linear groups are unirational over perfect fields of definition. (See Rosenlicht [8]; also Chevalley [3].) Therefore, let $f: P^n \to \mathfrak{g}$ be a rational map from a projective space P^n generically onto $\mathfrak{g}$. Let $\epsilon \in \mathfrak{g}$ be the identity, and suppose that $\sigma \in \mathfrak{g}$ is given. By 2.1, we can find an open subset $\mathfrak{U}$ of $\mathfrak{g}$ such that every point in $\mathfrak{U}$ is a regular image of some point on P^n. Since $\mathfrak{g}$ is a homogeneous space, we can next find two points σ', ϵ' in $\mathfrak{U}$ and a biregular mapping g of $\mathfrak{g}$ onto itself such that $g(\sigma') = \sigma$, $g(\epsilon') = \epsilon$. Take the composition $g \circ f: P^n \to \mathfrak{g}$; then there are points s, e on P^n at which $g \circ f$ is regular and $g \circ f(s) = \sigma$, $g \circ f(e) = \epsilon$.

Now let $m: \mathfrak{g} \times V \to V$ be the morphism defining the operation of $\mathfrak{g}$ on V, so that $m(\gamma, x) = \gamma x$ for every $\gamma \in \mathfrak{g}$ and for every x on V. Take the composition $m \circ (g \circ f \times 1_V): P^n \times V \to V$ (where 1_V is the identity morphism of V); this rational mapping is regular at $s \times x$ and at $e \times x$ for every x on V. Therefore, if one takes the straight line L in P^n passing through s and e, the last rational mapping is regular along $L \times V$ and, hence, can be restricted to a rational mapping $h: L \times V \to V$ which is regular at $s \times x$ and at $e \times x$ for every x on V.

In order to prove that $D = D^\epsilon$ and D^σ are linearly equivalent, we may assume that D is a variety. That being assumed, let U be the graph of the restriction of h to $L \times D$; if t, z are independent generic points over a field $K \supseteq k(s, e)$ on L, D, respectively, then U is the K-locus of the point $t \times z \times h(t, z)$ and is a subvariety of $L \times D \times V$. Consider now the intersection $U \cdot (s \times D \times V)$. Since the restriction of h to $L \times D$ is regular at each point of $s \times D$, the intersection has one and only one component of multiplicity one which is the K-locus of $s \times z \times \sigma z$. Take the algebraic projection $\mathrm{pr}_{L\times V}[U \cdot (s \times D \times V)]$. It is a cycle of the type $d \cdot (s \times D^{\sigma})$ with $d \geqq 0$ an integer. But $K(z) \subseteq K(\sigma, \sigma z) \subseteq K(s, \sigma z) = K(\sigma z)$; so that $d = 1$. Finally observe that $\mathrm{pr}_{L\times V}[U \cdot (s \times D \times V)] = (\mathrm{pr}_{L\times V} U) \cdot (s \times V)$ since D is complete and by [10, Th. 8, Chap. VIII, § 3, p. 229]. Thus we have established that $(\mathrm{pr}_{L\times V} U) \cdot (s \times V) = s \times D^{\sigma}$. A parallel argument will show that $(\mathrm{pr}_{L\times V} U) \cdot (e \times V) = e \times D^{\epsilon}$. It follows then that a correspondence $W = \mathrm{pr}_{L\times V} U$ exists between L and V such that $W(s) = D^{\sigma}$ and $W(e) = D^{\epsilon} = D$. Therefore [10, p. 328 ff.], D^{σ} is linearly equivalent to D.

Q. E. D.

2.3. Corollary. *Let V be a complete normal variety on which an algebraic linear group $\mathfrak{g}$ operates regularly. Let D be a divisor on V. Then, for each $\sigma \in \mathfrak{g}$ and $\tau \in \mathfrak{g}$, D^{σ} is linearly equivalent to D^{τ} whenever σ and τ belong to one and the same connected component.*

The proof of the corollary is essentially identical with that of 2.2 and is therefore omitted.

2.4. Corollary. *Let V, $\mathfrak{g}$, D be as in 2.3. Let $\mathfrak{g}_0$ be the connected component of the identity $\epsilon \in \mathfrak{g}$, and $\{\epsilon, \sigma, \cdots, \tau\}$ a complete set of representatives of the factor group $\mathfrak{g}/\mathfrak{g}_0$. Set $E = D^{\epsilon} + D^{\sigma} + \cdots + D^{\tau}$. Then, for every $\gamma \in \mathfrak{g}$, E^{γ} is linearly equivalent to E.*

Proof. For each $\gamma \in \mathfrak{g}$, the set $\{\gamma\epsilon, \gamma\sigma, \cdots, \gamma\tau\}$ is another complete set of representatives of $\mathfrak{g}/\mathfrak{g}_0$. Therefore, to every $\rho \in \{\epsilon, \sigma, \cdots, \tau\}$ there corresponds one and only one $\gamma\rho' \in \{\gamma\epsilon, \gamma\sigma, \cdots, \gamma\tau\}$ belonging to the same coset of $\mathfrak{g}_0$ as ρ. By 2.3, D^{ρ} and $D^{\gamma\rho'}$ are linearly equivalent; therefore, E and $E^{\gamma} = D^{\gamma\epsilon} + D^{\gamma\sigma} + \cdots + D^{\gamma\tau}$ are linearly equivalent. Q. E. D.

Let V be a complete normal variety, and D a divisor on V. The complete linear system $\mathscr{L}(D)$ is the set of positive divisors $\{D + \mathrm{div}(f); f \in \mathrm{L}(D)\}$. We say that D *separates points* of V if the linear system $\mathscr{L}(D)$ separates points of V, *viz.* if to each pair of distinct points x, y on V there is a divisor $X \in \mathscr{L}(D)$ such that the support of X contains x and not y. Let $\{f_0, f_1, \cdots, f_n\}$ be a basis of the vector space $L(D)$, and λ_D the mapping of V into P^n defined by $\lambda_D(u) = (f_0(u), \cdots, f_n(u))$, where u is a generic point on V over some common field of definition for the f_i (cf. Weil [10, Chap. IX, § 5]). As is well known, if λ_D' is another mapping of V into P^n defined similarly by a different choice of basis for $L(D)$, then $\lambda_D' = \alpha \circ \lambda_D$ holds true with a linear automorphism α of P^n. It is also known that λ_D is a set-theoretically injective morphism of V into P^n if and only if D

separates points [10, Th. 12, Chap. IX, § 5, p. 288]; when that is so, λ_D is a purely inseparable mapping from V onto its geometric image in P^n, and the latter coincides with the set-theoretical image of V because V is complete.

We are now in a position to state and prove our main results.

2.5. THEOREM. *Let V be a complete normal variety on which an algebraic linear group $\mathfrak{g}$ operates regularly. Then, there exist a birational mapping λ from V into a projective space P^n and a rational homomorphism $T: \mathfrak{g} \to PL(n)$, such that $T(\sigma)\lambda(x) = \lambda(\sigma x)$ for every $\sigma \in \mathfrak{g}$ and for every x generic on V over $k(\sigma)$. Moreover, the kernel of T consists of precisely those elements of $\mathfrak{g}$ which act as identity on V.*

Proof. Take an arbitrary birational mapping $f: V \to P^n$, and pull back the hyperplanes in P^n by f^{-1} [10, Th. 11, Chap. IX, § 5, p. 284], thereby obtaining a reduced linear system on V. Call D any one of the positive divisors in the linear system; obviously D is simple. With the notations of 2.4, let $E = D^\epsilon + D^\sigma + \cdots + D^\tau$. Then, E is again simple, because E is positive and satisfies $L(D) \subseteq L(E)$. Let T_E be as in 1.1 and 1.2, and set $\lambda = \lambda_E$ where λ_E is as in the remarks preceding the present theorem. If σ, x are given as above, it is easy to show that ${}^t(T_E(\sigma))\lambda(x) = \lambda(\sigma^{-1}x)$, where ${}^tM$ denotes the transposed of matrix M. Indeed, let $\{f_0, \cdots, f_n\}$ be a basis of $L(E)$, and let $f_i^\sigma g_\sigma = s_{i0}f_0 + \cdots + s_{in}f_n$ for $0 \leqq i \leqq n$. Then, the matrix S representing $T_E(\sigma)$ with reference to the basis $\{f\}$ has $(s_{i0}, \cdots, s_{in})$ as its i-th column. Apply ${}^tS$ to $\lambda(x) = (f_0(x), \cdots, f_n(x)) \in P^n$; then, the i-th homogeneous coordinate of ${}^tS\lambda(x)$ is

$$s_{i0}f_0(x) + \cdots + s_{in}f_n(x) = f_i^\sigma(x)g_\sigma(x) = f_i(\sigma^{-1}x)g_\sigma(x).$$

Therefore,

$$\begin{aligned} {}^tS\lambda(x) &= (f_0(\sigma^{-1}x)g_\sigma(x), \cdots, f_n(\sigma^{-1}x)g_\sigma(x)) \\ &= (f_0(\sigma^{-1}x), \cdots, f_n(\sigma^{-1}x)) = \lambda(\sigma^{-1}x), \end{aligned}$$

which establishes ${}^t(T_E(\sigma))\lambda(x) = \lambda(\sigma^{-1}x)$. Let now $T: \mathfrak{g} \to PL(n)$ be the rational homomorphism given by $\sigma \to {}^t(T_E(\sigma^{-1}))$; it is clear by 1.2 that the kernel of T is as asserted, and by what we saw just now we have $T(\sigma)\lambda(x) = \lambda(\sigma x)$. Q. E. D.

2.6. THEOREM. *Let V be a complete normal variety which admits a biregular model in a projective space P^m. Let $\mathfrak{g}$ be an algebraic linear group operating regularly on V. Then, there exist a (biregular) isomorphism λ from V into a projective space P^n and a rational homomorphism $T: \mathfrak{g} \to PL(n)$,* *such that* $T(\sigma)\lambda(x)=\lambda(\sigma x)$ *for every* $\sigma\in\mathfrak{g}$ *and for every* x *on* V. *Moreover, the kernel of* T *consists of precisely those elements of* $\mathfrak{g}$ *which act as identity on* V.

Proof. Without losing generality, we may assume that V is a subvariety of a projective space P^m to begin with. Let D be a hyperplane section of V, and let $E=D^\epsilon+D^\sigma+\cdots+D^\tau$ with the notations of 2.4. Since each one of $D^\sigma,\cdots,D^\tau$ separates points of V, we know that $D^\sigma+\cdots+D^\tau$ separates points; on the other hand, $D^\epsilon=D$ is a hyperplane section. Therefore, $E=D^\epsilon+(D^\sigma+\cdots+D^\tau)$ is ample (cf. [10, p. 286] and [6, Th. 6.31, p. 166]). Hence, if we let $\lambda=\lambda_E\colon V\rightarrow P^n$, the mapping λ is an isomorphism. Next define $T\colon\mathfrak{g}\rightarrow PL(n)$ by $T(\sigma)={}^t(T_E(\sigma^{-1}))$ for every $\sigma\in\mathfrak{g}$. With respect to λ and T, our assertions are readily verified by the same reasoning as in the proof of 2.5. This completes the proof.

INDIANA UNIVERSITY,
BLOOMINGTON, INDIANA.

REFERENCES.

[1] I. Barsotti, *Risultati e problemi nella teoria delle varietà gruppali*, Messina-Seminario Matematico dell'Università, 1959.

[2] A. Borel, "Groupes linéaires algébriques," *Annals of Mathematics*, Series 2, vol. 64 (1956).

[3] C. Chevalley, "On algebraic group varieties," *Journal of the Mathematical Society of Japan*, vol. 6 (1954).

[4] H. Matsumura, "On algebraic groups of birational transformations," *Rendiconti dell'Accademia Nazionale dei Lincei*, ser. 8, vol. 34 (1963).

[5] T. Matsusaka, "Polarized varieties, fields of moduli and . . .," *American Journal of Mathematics*, vol. 80 (1958).

[6] M. Nagata-Y. Nakai, *Algebraic Geometry* (in Japanese), Kyoritsu Shuppan, Tokyo, 1960.

[7] M. Rosenlicht, "Some basic theorems on algebraic groups," *American Journal of Mathematics*, vol. 78 (1956).

[8] ———, "Some rationality questions on algebraic groups," *Annali di Matematica*, ser. 4, vol. 43 (1957).

[9] A. Weil, "On algebraic groups of transformations," *American Journal of Mathematics*, vol. 77 (1955).

[10] ———, *Foundations of Algebraic Geometry* (*Revised and Enlarged*), American Mathematical Society, Providence, 1962.

A PROOF OF THE CONVERGENCE OF THE POINCARÉ-VON ZEIPEL PROCEDURE IN CELESTIAL MECHANICS.

By Richard Barrar.

(In Memory of Richard C. Kao)

Introduction. In 1954 A. N. Kolmogorov [5, 6] announced some theorems on the conservation of conditionally periodic motions with a small change of the Hamiltonian. He gave only a brief sketch of a proof. In an attempt to give a complete proof, J. Moser [7] was led in 1962 to his basic paper "On Invariant Curves of Area Preserving Mappings of an Annulus."

The present paper is an extension of the Moser procedure to give a proof of the Kolmogorov theorems. Strong use will be made of some canonical transformations introduced by Poincaré [8, Vol. II], to deal with perturbation problems in celestial mechanics. Later these transformations were also used by von Zeipel [11]. Interestingly, Poincaré considered these transformations, which introduce the so-called "small divisors," to be only asymptotic series that always diverge. Chapter XIII of Poincaré [8, Vol. II] is devoted to just this question.

Poincaré was correct in assuming that these series diverge at an everywhere dense set of frequencies. However, it was Kolmogorov's idea (going back to a paper of C. L. Siegel [10]) that there still could be a set of frequencies of density close to 1 where the series converge.

Just as this proof was completed, the author found that Arnol'd [1] had written a similar proof of the Kolmogorov conjecture. While Arnol'd's paper applies to n degrees of freedom, the present paper, for purposes of clear exposition, applies to only two degrees of freedom. The present paper differs from that of Arnol'd's in that it treats the important class of problems when the unperturbed Hamiltonian does not depend on both momenta. It is further felt that the technical details of the present paper are less formidable than those of Arnol'd's.

Main Theorem. We state and prove Theorem 1 for two degrees of freedom; however, there is no difficulty in extending it to n degrees of freedom. We further restrict ourselves to analytic Hamiltonians, although the work of Moser [7] shows that this restriction may be relaxed.

Received January 8, 1965.

THEOREM 1. *Consider a Hamiltonian of the form:*

$$(1)\qquad H = H_0(p_1) + \mu[H_1(p_1, p_2) + H_2(p_1, p_2, q_1, q_2)] \text{ with } 0 < \mu \leqq 1,$$

and corresponding canonical differential equations

$$(2)\qquad \frac{dp_i}{dt} = \frac{\partial H}{\partial q_i} \qquad \frac{dq_i}{dt} = -\frac{\partial H}{\partial p_i} \qquad i = 1, 2,$$

where all functions are assumed analytic, and H_2 is periodic of period 2π in q_1 and q_2, and of mean value zero.

For two fixed values $p_1{}^0$, $p_2{}^0$, let

$$\lambda_1 = \left.\frac{\partial H_0(p_1) + \mu H_1(p_1, p_2)}{\partial p_1}\right|_{p_i{}^0} \qquad \lambda_2 = \left.\frac{\partial H_1(p_1, p_2)}{\partial p_2}\right|_{p_i{}^0}$$

satisfy

$$\text{(H1)}\qquad \left| n_1 + n_2\frac{\mu\lambda_2}{\lambda_1} \right| \geqq \frac{\delta\mu}{(n_2)^2} \text{ for all integers } n_1, n_2, \text{ with } n_2 \neq 0.^1$$

Further, for some neighborhood of μ_0, with $0 \leqq \mu_0 \leqq 1$, let

$$\text{(H2)}\qquad \det \left| \frac{\partial^2(H_0(p_1) + \mu H_1(p_1, p_2))}{\partial p_i \partial p_j} \right|_{p_i{}^0} = 2N\mu \neq 0.$$

Then for $|H_2| \leqq \epsilon(\delta, N)$, and for μ in the above neighborhood, there is a conditionally periodic solution of (1) of the form:

$$\begin{aligned} q_1 &= \lambda_1(t - \tau_1) + \phi_1(e^{i\lambda_1 t}, e^{i\mu\lambda_2 t}) \\ q_2 &= \mu\lambda_2(t - \tau_2) + \phi_2(e^{i\lambda_1 t}, e^{i\mu\lambda_2 t}) \\ p_i &= A_i + \psi_i(e^{i\lambda_1 t}, e^{i\mu\lambda_2 t})\,;\ A_1, A_2 \text{ constants.} \end{aligned}$$

Outline of Proof of Theorem 1. It will be more convenient to write the Hamiltonian in the form:

$$(3)\qquad H = (\lambda_1 + \alpha_1)p_1 + (\mu\lambda_2 + \alpha_2)p_2 + F_0(p) + F_1(p, q) + \text{Constant}; \qquad 0 < \mu \leqq 1$$

with α_1, α_2 small constants, $F_0(p)$ a power series starting with terms of the second degree, and the mean value of $F_1(p, q)$ noted by $\bar{F}_1(p)$ being zero, i. e.,

$$\bar{F}_1(p) = \frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi} F_1(p, q)\, dq_1 dq_2 = 0.$$

---

[1] For $0 \leqq \omega \leqq M$, the exterior measure of all ω not satisfying $|n_1 + n_2\mu\omega| \geqq \frac{\delta\mu}{(n_2)^2}$ for all $n_2 \neq 0$ is less than $\frac{4\pi^2}{3}M\mu\delta$. (See C. L. Siegel [9], and Koskma [4, especially Kap. III, Kap V]).

This can clearly be done by making an expansion about the point $p_1{}^0$, $p_2{}^0$ of (H1), (H2) of Theorem 1, and then introducing new canonical variables $p_i' = p_i + p_i{}^0$, $q_i' = q_i$. In transforming from (1) to (3), we can make $\alpha_1 = \alpha_2 = 0$.

We begin by considering the Hamiltonian defined in the region

$$(3') \qquad N_0: \quad \begin{array}{c} |p_i| < R_0 \\ 1 - \rho_0 < |z_i| < 1 + \rho_0 \end{array}$$

where $z_j = e^{iq_j}$. [Since $F_1(p,q)$ is periodic of period 2π in the q_i, and is defined in a strip of the real q_i axis, the transformation to the z_i variables in an annulus is well defined; for many questions, it is more convenient to think of it as a function of the z_i].

We now proceed, by using a variant of the Poincaré-von Zeipel transformations, to define an infinite sequence of canonical transformations, the purpose being to eliminate the term $F_1(p,q)$ of (3). The motivation for the use of these canonical transformations is found in Poincaré [8, Vol. II]. In paragraph 125 Poincaré introduces the canonical transformation that we use in Lemma 1 (see below), and in paragraph 128, he introduces the translation that we use in Lemma 2 (see below). Von Zeipel [11] used similar transformations in his work in celestial mechanics, and in the current literature on celestial mechanics they are referred to as the von Zeipel transformations.

By the Poincaré transformations of Lemma 1 and 2, we will define a sequence of canonical transformations; at the n-th stage, we will transform the equation (3) into an equivalent equation in terms of the canonical variables $P_i{}^{(n)'}$, $Z_i{}^{(n)'}$, which will be defined in a region:

$$N_n: \quad \begin{array}{c} P_i{}^{(n)} = P_i{}^{(n)'} + A_i{}^{(n)}, \; |P_i{}^{(n)'}| < R_n \\ 1 - \rho_n < |Z_i{}^{(n)'}| < 1 + \rho_n \end{array}$$

with the notation to be used throughout the paper:

$$R_{n+1} = (R_n)^{(4/3)}, h_{n+1} = (h_n)^{4/3}, R_n = (h_n)^{15}, Z_j{}^{(n)'} = Z_j{}^{(n)} = e^{iQ_j{}^{(n)}}.$$

$$|\xi_i{}^{(j)}| < R_i(\xi_i{}^{(j)} \text{ a constant}), \; A_j{}^{(n)} = \sum_{i=1}^{n} \xi_i{}^{(j)},$$

$$\tilde{h}_n = 2\sum_{j=0}^{n-1} h_j, \; h_0 = h, \; \rho_n = \rho_0 - \tilde{h}_n.$$

The Hamiltonian in terms of the new variables $P^{(n)'}$, $Q^{(n)'}$ is still of the form (3), but with the corresponding α_1, α_2 and $F_1(p,q)$ all approaching zero; and p, q are analytic functions of $P^{(n)'}$, $Q^{(n)'}$.

The proof is completed by showing that the original variables p_1, q_1, p_2, q_2 can be expressed in terms of A_1, A_2, $Q_1{}^{\infty}$, $Q_2{}^{\infty}$ where $A_j = \sum_{i=1}^{\infty} \xi_i{}^{(j)}$, and $Q_1{}^{(n)} \rightarrow Q_1{}^{\infty} = \lambda_1(t-\tau_1)$, $Q_2{}^{(n)} \rightarrow Q_2{}^{\infty} = \mu\lambda_2(t-\tau_2)$.

We now state and prove Lemmas 1 and 2, which are concerned with one particular canonical transformation of the infinite sequence.

LEMMA 1. *If in the region*

$$N_n: \begin{cases} |\,p_i\,| < R_n \\ 1-\rho_n < |\,z_i\,| < 1+\rho_n \end{cases}$$

$$F_1{}^{(n)}(p,q) = \sum{}' A_{n_1n_2}(p_1,p_2)\, e^{i(n_1q_1+n_2q_2)} \quad \textit{where } \textstyle\sum{}' \textit{ means no terms with } n_1 = n_2 = 0$$

$$= \sum{}' A_{n_1n_2}(p_1,p_2)\,(z_1)^{n_1}(z_2)^{n_2}$$

with

$$|\,F_1{}^{(n)}(p,q)\,| \leqq \mu R_n{}^{8/3} \qquad\qquad 0 < \mu < 1$$

then the generating function

$$W_{n+1} = P_1q_1 + P_2q_2 + S(P,q)$$

where

$$\text{(H1)} \qquad \lambda_1 \frac{\partial S(P,q)}{\partial q_1} + \mu\lambda_2 \frac{\partial S(P,q)}{\partial q_2} = -F_1{}^{(n)}(P,q)$$

or

$$S(P,q) = -\sum{}' \frac{A_{n_1n_2}(P_1,P_2)\,(z_1)^{n_1}(z_2)^{n_2}}{n_1\lambda_1 + n_2\mu\lambda_2}$$

[*and where we further assume* λ_1, $\mu\lambda_2$ *satisfy* (H2) *of Theorem* 1 *with* $\delta \geqq h_n$] *defines a canonical transformation between the variables* (p,q) *and* (P,Q) *by means of the formulas*:

$$\text{(4)} \qquad \text{(a)} \quad Q_i = q_i + \frac{\partial S(P,q)}{\partial P_i} \qquad\qquad \text{(b)} \quad p_i = P_i + \frac{\partial S(P,q)}{\partial q_i}.$$

This transformation may further be put into the form:

$$\text{(5)} \qquad \text{(a)} \quad q_i = Q_i + L_i(P_i, Q_i) \qquad\qquad \text{(b)} \quad p_i = P_i + K_i(P,Q)$$

where the function K_i, L_i *are defined and analytic at least in the region*

$$\bar{N}_n:\ |\,P_i\,| \leqq 1/2\, R_n \qquad 1-(\rho_n - 2h_n) < |\,Z_i\,| < 1 + (\rho_n - 2h_n)$$

and in this region:

(6) (a) $|L_i| \leqq h_n$ (b) $|K_i| \leqq (1/2)R_{n+1} < (1/4)R_n$

thus assuring (5) *maps the* P, Q *region* $\tilde{N}_n$ *into the* p, q *region* N_n.

Proof. To simplify the notation we give the proof only in the case $n = 0$, and set $F_1^{(0)} = F_1$. Since the Cauchy integral formula applies for functions of several complex variables to polycylinders (See Bochner and Martin [2, pp. 31-32], or H. Cartan [3, Chapter IV, Par. 5]), it readily follows that the uniquely determined Laurent series for $F_1(P, z)$ is majorized term-by-term by the function $\mu M(P, z)$ with

$$M(P,z) = R_0^{8/3} \prod_{i=1}^{2} \left[\frac{1+\rho_0}{((1+\rho_0)-z_i)} - \frac{(1-\rho_0)}{((1-\rho_0)-z_i)}\right]\left[\frac{R_0}{(R_0-P_i)}\right]$$

$$= R_0^{8/3} \prod_{i=1}^{2} \left[\sum_{n=0}^{\infty}\left(\frac{z_i}{1+\rho_0}\right)^n + \sum_{n=1}^{\infty}\left(\frac{1-\rho_0}{z_i}\right)^n\right]\left[\sum_{n=0}^{\infty}\left(\frac{P_i}{R_0}\right)^n\right].$$

We write this property that $\mu M(P, z)$ majorizes $F_1(P, z)$ as

$$F_1(P,z) \ll \mu M(P,z).$$

Since λ_1, $\mu\lambda_2$ satisfy (H2) of Theorem 1, with $\delta > h$, we further have:

$$S(P,q) \ll \left[\frac{1}{h}\left(z_2 \frac{\partial}{\partial z_2}\right)^2 + \frac{\mu}{\lambda_1}\right] M(P,z)$$

thus assuring the existence of $S(P, q)$. Furthermore it readily follows that

$$\text{(7a)} \qquad \frac{\partial S(P,q)}{\partial q_i} \ll z_i \frac{\partial}{\partial z_i}\left[\frac{1}{h}\left(z_2 \frac{\partial}{\partial z_2}\right)^2 + \frac{\mu}{\lambda_1}\right] M(P,z)$$

$$\text{(7b)} \qquad \frac{\partial S(P,q)}{\partial P_i} \ll \frac{\partial}{\partial P_i}\left[\frac{1}{h}\left(z_2 \frac{\partial}{\partial z_2}\right)^2 + \frac{\mu}{\lambda_1}\right] M(P,z).$$

Thus for $|P_i| < (1/2)R_0$ and $1 - (\rho_0 - h) < |z_i| < 1 + (\rho_0 - h)$

$$\text{(8a)} \qquad \left|\frac{\partial S(P,q)}{\partial P_i}\right| \leqq \frac{CR_0^{8/3}}{h^4 R_0} < \frac{h}{2}.$$

In (8a) we use the convention that C is some number that does not depend on h or μ. Thus the last inequality will hold if h is chosen small enough, and then it will surely hold for every h_n (since $h_n < h$).

We wish first to use this estimate to discuss the invertibility of the transformation

$$\text{(4a)}' \qquad Q_j = q_j + \frac{\partial S}{\partial P_j}(P,z)$$

or the equivalent transformation

$$(4a)'' \qquad Z_j = z_j e^{i\frac{\partial S}{\partial P_j}} = z_j + (e^{i\frac{\partial S}{\partial P_j}} - 1) z_j.$$

It is clearly equivalent for if we solve for $z_i = \tilde{L}_i(P, Z)$ in (4a)″, we can substitute this z_i in equation (4a)′. We prefer to deal with the equation in the form (4a)″. Using the power series for e^x, it readily follows from the estimate (8a) that (4a)″ is equivalent to

$$(9)_j \qquad Z_j = z_j + f_j(P, z)\,;\, |f_j(P, z)| \leqq 3/4\,h \qquad j = 1, 2$$

in

$$|P_i| < (1/2) R_0 ; 1 - (\rho_0 - h) < |z_j| < 1 + (\rho_0 - h)$$

(For brevity let us call this the annulus $\rho_0 - h$.)

In the equation $(9)_2$, consider P_1, P_2, z_1 as fixed, then the equation is an equation between the two complex variables Z_2, z_2. By Rouché's theorem, applied to z_2 in the annulus $\rho_0 - h$, every point Z_2 in the annulus $\rho_0 - 2h$ is the image of one and only one point z_2. The Jacobian of the transformation cannot vanish, otherwise the transformation would not be locally one to one. Hence by the implicit function theorem for analytic functions,

$$(10) \qquad z_2 = Z_2 + g_2(P_1, P_2, z_1, z_2).$$

Further by applying Rouché's theorem to a circle of radius h around Z_2, it follows that $|g_2| \leqq h$.

We may now substitute (10) in equation $(9)_1$ to obtain an equation between Z_1 and z_1 for fixed Z_2, P_1, P_2. By applying Rouché's theorem again, we obtain (5a), and (6a), first for z_1; then by using (10) we obtain them for z_2.

The second transformation equation (5b) follows from (4b) and (5a). (6b) follows from the estimate resulting from (7a):

$$(8b) \qquad \left|\frac{\partial S}{\partial q_i}(P, q)\right| \leqq \frac{CR_0^{8/3}}{h^5} < \tfrac{1}{2} R_1 < \tfrac{1}{4} R_0$$

$$1 - (\rho_0 - h) < |Z_i| < 1 + (\rho_0 - h)$$
$$\text{and } |P_i| < R_0.$$

Lemma 2. *Assume in the region*

$$N_n: \begin{cases} |p_i| < R_n \\ 1 - \rho_n < |z_i| < 1 + \rho_n \end{cases}$$

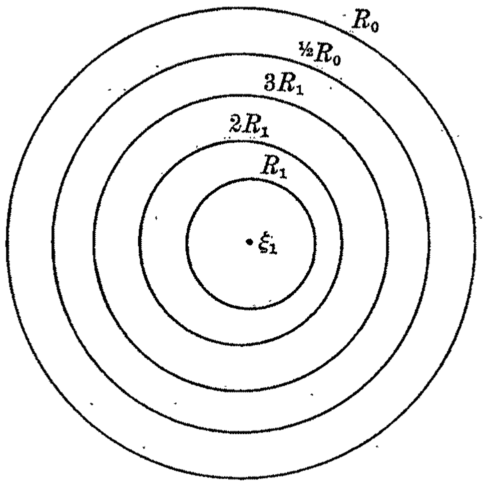


FIGURE 1. Schematic Representation of p Space.

the Hamiltonian has the form:

$$H = (\lambda_1 + \alpha_1^{(n)})p_1 + (\mu\lambda_2 + \alpha_2^{(n)})p_2 + F_0^{(n)}(p) + F_1^{(n)}(p,q) + constant$$

with $F_0^{(n)}$ *starting with terms of the second degree; and the mean value of* $F_1^{(n)}(p,q)$ *being zero; and with* λ_1, $\mu\lambda_2$ *satisfying* (H2) *of Theorem* 1 *with* $\delta > h_n$; *and* $\alpha_i^{(n)}$ *constants; and the* 4 *following hypotheses hold:*

(H1) $|F_1^{(n)}(p,q)| \leqq \mu R_n^{8/3}$

(H2) $|\alpha_i^{(n)}| \leqq \mu R_n^{5/3} \quad i = 1, 2$

(H3) $\det \left| \dfrac{\partial^2 F_0^{(n)}}{\partial p_i \partial p_j} \right| \geqq \mu N > 0$ in R_n

(H4) $\dfrac{\partial^2 F_0^{(n)}}{\partial p_1^2}$ *and* $\dfrac{\partial^3 F_0^{(n)}}{\partial p_1^3}$ *are bounded by* M, *and all other second and third partials of* $F_0^{(n)}$ *are bounded by* μM, *in* R_n.

Then by introducing the canonical transformation W_{n+1} *of Lemma* 1, *and a further translation*

$$P_i' = P_i - \xi_i \qquad Q_i' = Q_i$$

with suitably chosen ξ_i *in* R_{n+1}, *the new Hamiltonian in terms of* P_i', Q_i' *in the region:* (For the regions R_0 and R_1 see Figure 1.)

$$N_{n+1}: \begin{cases} |P_i'| < R_{n+1} \\ 1 - \rho_{n+1} < |Z_i'| < 1 + \rho_{n+1} \end{cases}$$

will have the form:

$$H = (\lambda_1 + \alpha_1^{(n+1)})P_1' + (\mu\lambda_2 + \alpha_2^{(n+1)})P_2' + F_0^{(n+1)}(P') + F_1^{(n+1)}(P', Q') + \textit{constant}$$

$$F_0^{(n+1)}(P') = F_0^{(n)}(P') + D^{(n+1)}(P');$$

with $F_0^{(n+1)}(P')$ *starting with terms of the second degree, and the mean value of* $F_1^{(n+1)}(P', Q')$ *being zero, and the* 3 *following conclusions hold:*

C1. $|F_1^{(n+1)}(P', Q')| \leqq \mu R_{n+1}{}^{8/3}$

C2. $|\alpha_i^{(n+1)}| \leqq \mu R_{n+1}{}^{8/3} \qquad i = 1, 2$

C3. $\left|\dfrac{\partial^2 D^{(n+1)}}{\partial P_i \partial P_j}\right|$ *and* $\left|\dfrac{\partial^3 D^{(n+1)}}{\partial P_i \partial P_j \partial P_k}\right| \leqq \mu h_n{}^8, R_n = (h_n)^{15}, R_{n+1} = (R_n)^{4/3}$.

Proof. To simplify the notation we will again give the proof only for the case $n = 0$, and we will not indicate the superscript (0).

Using (4b) the Hamiltonian H may be expressed as:

$$H = (\lambda_1 + \alpha_1)(P_1 + \frac{\partial S}{\partial q_1}(P, q)) + (\mu\lambda_2 + \alpha_2)(P_2 + \frac{\partial S}{\partial q_2}(P, q)) + F_0(P_i + \frac{\partial S}{\partial q_i}(P, q)) + F_1(P_i + \frac{\partial S}{\partial q_i}, q_i)$$

$$\begin{aligned} H = {} & (\lambda_1 + \alpha_1)P_1 + (\mu\lambda_2 + \alpha_2)P_2 + F_0(P) \\ & + \lambda_1 \frac{\partial S}{\partial q_1}(P, q) + \mu\lambda_2 \frac{\partial S}{\partial q_2}(P, q) + F_1(P, q) \\ & + D_1 + D_2 + D_3 \end{aligned}$$

with

$$D_1 = F_1(P + \frac{\partial S}{\partial q}, q) - F_1(P, q) = \sum_{i=1}^{2} \frac{\partial F_1}{\partial P_i}(P + \theta \frac{\partial S}{\partial q}, q) \frac{\partial S}{\partial q_i}$$

$$D_2 = \sum_{i=1}^{2} \frac{\partial F_0}{\partial P_i}(P) \frac{\partial S}{\partial q_i} + \alpha_i \frac{\partial S}{\partial q_i}$$

$$D_3 = \sum_{\substack{i=1 \\ j=1}}^{2} \frac{\partial^2 F_0}{\partial P_i \partial P_j}\left(P + \theta \frac{\partial S}{\partial q}\right)\left(\frac{\partial S}{\partial q_i}\right)\left(\frac{\partial S}{\partial q_j}\right).$$

The second line in the expression above for H vanishes, by our construction of the transformation S. (See (H1) of Lemma 1).

Let

$$D = D_1 + D_2 + D_3.$$

We next wish to show that in the region $|P_i| < 2R_1$, $1-(\rho_0-2h) < |Z_i| < 1+(\rho_0-2h)$, the remaining terms have the following bounds:

$$(11) \qquad |D| < \mu R_1^{8/3}$$

and if we define $\bar{D}$ by

$$\bar{D} = \frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi}(D_1+D_2+D_3)\,dQ_1dQ_2$$

then

$$(12) \qquad \left|\frac{\partial\bar{D}}{\partial P_i}\right| < \mu\,\frac{R_1^{8/3}}{R_0} < \mu R_1^{5/3}$$

$$(13) \qquad \left|\frac{\partial^2\bar{D}}{\partial P_i\partial P_j}\right| < \frac{\mu R_1^{8/3}}{R_0^2} < \mu R_0 h^8 < \mu h^8$$

$$(14) \qquad \left|\frac{\partial^3\bar{D}}{\partial P_i\partial P_j\partial P_k}\right| < \frac{\mu R_1^{8/3}}{R_0^3} = \mu R_0^{5/9} < \mu h^8.$$

We begin by estimating the term D_1. First we note that for $|P_i| \leqq (1/2)R_0$ it follows from (8b) that $|P_i+\theta\dfrac{\partial S}{\partial q_i}| \leqq \frac{3}{4}R_0$. Thus for $|P_i| \leqq \frac{1}{2}R_0$, one obtains the estimate:

$$|D_1| \leqq \mu C\,\frac{(R_0)^{8/3}}{R_0}\cdot\frac{(R_0)^{8/8}}{h^5} = C\mu R_0^4 = C\mu R_1^3 < \frac{\mu}{3}R_1^{8/3}$$

from (8b) and from analogous estimates for $|\partial F_1/\partial P|$. Since this estimate holds for $|P_i| = (\frac{1}{2})R_0$, and

$$|\tfrac{1}{2}R_0 - 2R_1| = \tfrac{1}{2}R_0\,|1-4R_0^{1/3}| > \tfrac{1}{4}R_0 \text{ for } h < \tfrac{1}{2}$$

the estimates analogous to (12), (13), and (14) for the partial derivative of $\bar{D}_1$ readily follow from the Cauchy integral formula.

To make the estimates for D_3 we first note that three of the four second partial derivatives of F_0 are bounded by μM, while $\dfrac{\partial^2F_0}{\partial P_1^2}$ is only bounded by M for $|P_i| \leqq \frac{1}{2}R_0$ and $|P_i+\theta\dfrac{\partial S}{\partial q_i}| \leqq \frac{3}{4}R_0$. However it follows from the partial differential equation which S satisfies (See (H1) of Lemma 1), that since F_1 and $\dfrac{\mu\partial S}{\partial q_2}$ are bounded by $\dfrac{\mu CR_0^{8/3}}{h^5}$ that

$$\left|\frac{\partial S}{\partial q_1}\right| \leqq \mu C\,\frac{R_0^{8/3}}{h^5}.$$

Thus we readily establish:

$$|D_3| \leqq \mu C \left(\frac{R_0^{8/3}}{h^5}\right)^2 = \mu \frac{CR_1^4}{h^{10}} \leqq \frac{\mu R_1^{8/3}}{3}.$$

The estimates for the partial derivatives of $\bar{D}_3$ corresponding to (12), (13), and (14) follows as above from Cauchy's formula.

The estimate for D_2 proceeds as follows: First, noting that F_0 starts with second degree terms, one estimates by the mean value theorem, and by reasoning as used in the estimate for D_3

$$\left| \sum_{i=1}^{2} \frac{\partial F_0}{\partial P_i} \frac{\partial S}{\partial q_i} \right| \leqq \sum_{i,j=1}^{2} \left| \frac{\partial^2 F_0}{\partial P_i \partial P_j}(\theta P) \right| |P_j| \left| \frac{\partial S}{\partial q_1} \right|$$
$$\leqq \mu C |P_j| \frac{R_0^{8/3}}{h^5},$$

or for $|P_i| \leqq 3R_1$

$$|D_2| \leqq CR_1 \frac{R_0^{8/3}}{h^5} \leqq \frac{CR_1^3}{h^5} = \frac{CR_1^{1/3} R_1^{8/3}}{h^5} \leqq \tfrac{1}{3} R_1^{8/3}.$$

To now obtain bounds for the partial derivatives of $\bar{D}_2$ from Cauchy's formula in $2R_1$ we would have to divide by R_1, not $\frac{1}{4}R_0$ as above; this will not result in estimates analogous to (12), (13), and (14). To remedy this we must give a more accurate estimate for $\bar{D}_2$ or what is equivalent $\partial \bar{S}/\partial q_i$. Since

$$\frac{\partial \bar{S}}{\partial q_2}(P_1, P_2) = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} \frac{\partial S}{\partial q_2}(P, q)\, dQ_1 dQ_2,$$

if we set $q_i = Q_i - \dfrac{\partial S}{\partial P_i}(P, q)$ in the integrand and expand about $q = Q$ by Taylor's formula, the first term in the integral will vanish because S is periodic, and we will obtain:

$$\left| \frac{\partial \bar{S}}{\partial q_2}(P_1, P_2) \right| \leqq \sum_{i=1}^{2} C \left| \frac{\partial^2 S}{\partial q_2 \partial q_i} \right| \left| \frac{\partial S}{\partial P_i} \right| \leqq \frac{C (R_0)^{8/3}}{h^6} \frac{R_0^{8/3}}{h^4 R_0} < 1/9 \frac{R_1^2}{h^5} R_0$$

where the bound follows from (8a) and analogous formulas for $\partial^2 S/\partial q_2 \partial q_i$.

From the partial differential equation (H1) of Lemma 1, one obtains a similar bound for $\partial \bar{S}/\partial q_1$, but with a factor μ. Using these estimates one now obtains, similar to our previous estimate for D_2:

$$|\bar{D}_2| < \frac{\mu}{3} R_1^{8/3} R_0 \text{ for } |P_i| \leqq 3R_1.$$

From this formula, the desired estimates analogous to (12), (13), and (14) for the partials of $\bar{D}_2$ for $|P_i| \leqq 2R_1$ now follow from Cauchy's formula. Thus (11), (12), (13), and (14) are established.

The proof will now be complete if we can show the existence of a point $\xi = (\xi_1, \xi_2)$ with $|\xi_i| < R_1$; where if $F_0^{(1)} = F_0 + D^{(1)}$

$$\frac{\partial (F_0^{(1)})}{\partial P_i}\bigg|_{P=\xi} = \alpha_i^{(1)} - \alpha_i \text{ with } |\alpha_i^{(1)}| < \mu R_1^{5/3}, \tag{15}$$

because the expansion of $F_0^{(1)}$ about ξ will then satisfy all the conclusions of Lemma 2.

We assert that the point ξ determined from the linear equations

$$\sum_{j=1}^{2} \frac{\partial^2 F_0^{(1)}}{\partial P_i \partial P_j}\bigg|_{P=0} \xi_j = -\alpha_i - \frac{\partial \bar{D}}{\partial P_i}\bigg|_{P=0} \tag{16}$$

is such a point.

For since α_i and $\partial \bar{D}/\partial P_i$ are both $\leqq \mu R_0^{5/3}$; and since the determinant is $\geqq \mu N$, it follows that ξ_1 and ξ_2 are $\leqq C R_0^{5/3}/N < R_1$. Further by writing out Cramer's formula, it follows that $|\xi_1| \leqq \mu R_1$. Application of Taylor's formula now shows:

$$\frac{\partial F_0^{(1)}}{\partial P_i}\bigg|_{P=\xi} = \frac{\partial \bar{D}}{\partial P_i}\bigg|_{P=0} + \sum_{j=1}^{2} \frac{\partial^2 F_0^{(1)}}{\partial P_i \partial P_j}\bigg|_{P=0} \xi_j + \frac{1}{2} \sum_{\substack{j=1 \\ k=1}}^{2} \frac{\partial^3 F_0^{(1)}}{\partial P_i \partial P_j \partial P_k}\bigg|_{P=\theta\xi} \xi_j \xi_k$$

which combined with (16) shows that $\alpha_i^{(1)}$ is the last term of this expression. Now using the bounds on the third order partial derivatives of $F_0^{(1)}$, together with $|\xi_1| \leqq \mu R_1$, $|\xi_2| \leqq R_1$; it follows that $|\alpha_i^{(1)}| \leqq \mu C R_1^2 \leqq \mu R_1^{5/3}$, which completes the proof.

Completion of proof by construction of infinite sequence of transformations. Lemmas 1 and 2 were concerned with one particular canonical transformation. We now use them to construct an infinite sequence of transformations. We assume the Hamiltonian is defined in the region N_0 (See 3′), and we also assume that $h > 0$, with

$$\rho_0 - 7h > 0, \tag{17}$$

and redefine R_0 if necessary so $R_0 = h^{15}$ (see Moser [7, page 5]). To construct the first transformation we must be assured that the hypotheses (H1)-(H4) of Lemma 2 are met. To meet the assumption H1 of Lemma 2, F_1 must be so small that in N_0:

$$|F_1| \leqq \mu R_0^{8/3};$$

Assumption (H2) of Lemma 2 is trivially met since to begin with $\alpha_1 = \alpha_2 = 0$. To satisfy (H3) and (H4) of Lemma 2, we first note that it follows from (H2) of Theorem 1 and from continuity that for small enough R_0; and for all $|p_i| \leqq R_0$:

$$\det\left|\frac{\partial^2 F_0}{\partial p_i \partial p_j} + b_{ij}\right| \geqq \mu N; \left|\frac{\partial^2 F_0}{\partial p_i \partial p_j} + b_{ij}\right| \leqq \begin{matrix} M \text{ if } i = j = 1 \\ \mu M \text{ otherwise} \end{matrix} \tag{18}$$

$$\left|\frac{\partial^3 F_0}{\partial p_i \partial p_j \partial p_k} + b_{ijk}\right| \leqq \begin{matrix} M \text{ if } i = j = k = 1 \\ \mu M \text{ otherwise} \end{matrix}$$

for arbitrary

$$|b_{ij}| \leqq 4\mu h$$

$$|b_{ijk}| \leqq 4\mu h.$$

Hence (H3) and (H4) of Lemma 2, are satisfied. With all assumptions now met the canonical transformation of Lemma 1, together with the translation specified in Lemma 2, takes the Hamiltonian into one defined in N_1, with corresponding:

$$|F_1^{(1)}| \leqq \mu R^{8/3} \qquad |\alpha_i^{(1)}| \leqq \mu R^{5/3}$$

and further it follows from the conclusions of Lemma 2, that all assumptions (see paragraph below) for mapping from N_2 into N_1 are met. Proceeding in this way we define an infinite sequence of mappings where now we start at the n-th stage in N_n, and we then obtain at the $n+1$ stage a Hamiltonian defined in N_{n+1} with corresponding:

$$|F_1^{(n+1)}| \leqq \mu R_{n+1}^{8/3} \qquad |\alpha_i^{(n+1)}| \leqq \mu R_{n+1}^{5/3}. \tag{19}$$

Note however that at each stage we add on a term $D^{(n)}(P)$ to $F_0^{(n-1)}$ as follows from the conclusions of Lemma 2. Thus the $F_0^{(n)}$ at the n-th stage is

$$F_0^{(n)} = F_0 + \sum_{j=1}^{n} D^{(j)}$$

but since

$$\frac{\partial^3 D^{(j)}}{\partial p_i \partial p_j \partial p_k} \text{ and } \frac{\partial^2 D^{(j)}}{\partial p_i \partial p_j} \leqq \mu (h_{j-1})^8$$

it readily follows from (18) that assumptions (H3) and (H4) will be fulfilled at each stage of the induction.

Since

$$\sum_{j=0}^{\infty} (h_j) < 3h, \tag{20}$$

the $F_1^{(n)}(p,q)$ uniformly approach zero in

$$(21) \qquad 1-(\rho_0-6h) < |Z_i| < 1+(\rho_0-6h)$$

which by (17) is non-vacuous.
Further from (17), (18), (19), (21) it follows that for $|P_i^{(n)\prime}| \leqq (\tfrac{1}{2})R_n$, and Q_i real that:

$$(22) \qquad \left|\frac{\partial H}{\partial P_i^{(n)\prime}}-\lambda_1\right| \leqq 3\mu R_n; \ \left|\frac{\partial H}{\partial Q_i}\right| \leqq \frac{\mu R_n^{8/3}}{h}$$

$$\left|\frac{\partial H}{\partial P_2^{(n)\prime}}-\mu\lambda_2\right| \leqq 3\mu R_n.$$

Hence it follows from the canonical differential equations (2) that

$$(23) \qquad P_i^{(n)} \to A_i = \sum_{j=1}^{\infty} \xi_j^i; \ Q_1^{(n)} \to \lambda_1(t-\tau_1) = Q_1^{(\infty)};$$
$$Q_2^{(n)} \to \mu\lambda_2(t-\tau_2) = Q_2^{\infty},$$

uniformly for any finite interval of time.

It only remains to show that we can express the original variables p_1, p_2, q_1, q_2 in terms of A_1, A_2, $Q_1^{(\infty)}$, $Q_2^{(\infty)}$.

We first observe that the regions N_n are nested

$$\begin{array}{ccccccccc} N_0 & \supset & N_1 & \supset & N_2 & \supset & N_3 & \supset & N_4 \cdots \\ & \leftarrow & & \leftarrow & & \leftarrow & & \leftarrow & \\ & W_1 & & W_2 & & W_3 & & W_4 & \end{array}$$

and that the transformation defined by Lemma 1, which we still call W_n, maps N_n into N_{n-1}.

Thus at the n-th stage

$$\binom{p}{q}\Bigl(= W_1 \circ W_2 \circ \cdots \circ W_n \begin{pmatrix} P^{(n)\prime}+A^n \\ Q^{(n)} \end{pmatrix} = \tilde{W}_n \begin{pmatrix} P^{(n)\prime}+A^n \\ Q^{(n)} \end{pmatrix}$$

or in terms of functions

$$p_i = P_i^{(n)\prime} + A_i^n + w_{npi}(P^{(n)\prime}+A^{(n)}, Q^{(n)})$$
$$q_i = Q_i^{(n)\prime} + w_{nqi}(P^{(n)\prime}+\mathrm{A}^{(n)}, Q^{(n)}).$$

Since the mappings $\tilde{W}_n$ all have range N_0; it readily follows that the analytic functions w_{npi} and w_{nqi} are bounded. In fact from repeated application of Lemma 1, it follows that this bound is $3h$ (see (20)). To show that they converge to unique analytic functions, it will be sufficient to show they form a uniform Cauchy sequence in some annulus, say, the annulus $(\rho_0 - 7h)$ which by (17) is non-vacuous.

We will estimate the difference $\tilde{W}_{n+1} - \tilde{W}_n$ below. (Since all differences $\tilde{W}_{n+p} - \tilde{W}_n$ can be written as sums of this form, the conclusion will readily follow.)

$$\begin{aligned}\tilde{W}_{n+1}\begin{pmatrix}P^{(n+1)'} + A^{n+1}\\ Q\end{pmatrix} - \tilde{W}_n\begin{pmatrix}P^{(n+1)'} + A^{n+1}\\ Q\end{pmatrix} &= \tilde{W}_n\left(W_{n+1}\begin{pmatrix}P^{(n+1)'} + A^{n+1}\\ Q\end{pmatrix}\right) - \tilde{W}_n\begin{pmatrix}P^{(n+1)'} + A^{n+1}\\ Q\end{pmatrix}\\ &= \tilde{W}_n\begin{pmatrix}P^{(n+1)'} + A^{n+1} + K_i\\ Q + L_i\end{pmatrix} - \tilde{W}_n\begin{pmatrix}P^{(n+1)'} + A^{n+1}\\ Q\end{pmatrix}\end{aligned}$$

where the K_i and L_i are the functions of Lemma 1, so $|L_i| \leqq h_n$; $|K_i| \leqq \frac{1}{2}R_{n+1}$. Thus in terms of functions, say for example w_{nqi}, we have:
If

$$\begin{aligned}d &= w_{(n+1)qi}(P^{(n+1)'} + A^{(n+1)}, Q) - w_{nqi}(P^{(n+1)'} + A^{(n+1)}, Q)\\ &= w_{nqi}(P^{(n+1)'} + A^{(n+1)} + K_i, Q + L_i) - W_{nqi}(P^{(n+1)'} + A^{(n+1)}, Q),\end{aligned}$$

then

$$|d| \leqq \left|\frac{\partial w_{nqi}}{\partial P}\right| |K_i| + \left|\frac{\partial w_{nqi}}{\partial Q}\right| |L_i| \leqq 3h\left[\frac{2}{R_n} R_{n+1} + \frac{1}{h} h_n\right] \leqq Chh_{n-1}$$

where the estimate for the partials of w_{nqi} follow from Cauchy's formula, together with $|R_n - 3R_{n+1}| > (\frac{1}{2})R_n$ which holds for $h < \frac{1}{2}$. Thus the w_{nqi} and similarly the w_{npi} form Cauchy sequences, and each converges to a unique analytic limit. This completes the proof. It further shows that the limiting functions are bounded by a constant times h.

SYSTEM DEVELOPMENT CORPORATION,
SANTA MONICA, CALIFORNIA.

REFERENCES.

[1] V. I. Arnol'd, "Proof of a theorem of A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian," *Russian Mathematical Surveys* Vol. 18, No. 5, 1963, pp. 9-36. Translation of *Uspekhi Mathematicheskikh Nauk*, Vol. 18.

[2] S. Bochner and W. T. Martin, *Several complex variables*, Princeton University Press (1948).

[3] H. Cartan, *Elementary theory of analytic functions of one or several complex variables*, Addison-Wesley (1963).

[4] J. F. Koksma, *Diophantische Approximationen*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. IV, No. 4, Springer-Verlag, 1936. Reprinted by Chelsea Publishing Company, New York.

[5] A. N. Kolmogorov, "General theory of dynamical systems and classical mechanics," *Proceedings of the International Congress of Mathematicians*, 1954, Vol. 1, pp. 315-333. Amsterdam, 1957.

[6] ———, "The conservation of conditionally periodic motions with a small change in the Hamiltonian," *Doklady Akademii Nauk* SSSR, vol. 98 (1954), pp. 527-530.

[7] J. Moser, "On invariant curves of area-preserving mappings of an annulus," *Nachrichten der Akademie Wissenschaften Gottingen, Mathematik-Physik Klasse*, vol. II(1962), pp. 1-20.

[8] H. Poincaré, *Les méthodes nouvelles de la méchanique celeste*, Vol. II, Dover Reprint, 1957.

[9] C. L. Siegel, *Vorlesungen Über Himmelsmechanik*, Sections 23, 24, Springer-Verlag, 1956.

[10] ———, "Iteration of analytic functions," *Annals of Mathematics*, vol. 43 (1942), pp. 607-612.

[11] H. von Ziepel, "Recherches sur le mouvement des petites planèts," *Arkiv foer Matematik, Astronomi och Fysik*, vol. 11, No. 1 (1916), pp. 1-58.

ON THE GRADED RING OF THETA-CONSTANTS (II).[1]

By Jun-ichi Igusa.

Introduction. In this paper, we shall remark that our "fundamental lemma" on theta-constants [3] can be generalized in such a way that we shall be able to investigate the ring of Siegel modular forms of an arbitrary level with the aid of theta-constants. In doing so, we shall make a few simplifications over the previous treatment. We shall be concerned about theta-constants for the principal matrix

$$E=\begin{bmatrix} 0 & 1_g \\ -1_g & 0 \end{bmatrix}$$

and about subgroups of the Siegel modular group $Sp(g,\boldsymbol{Z})$. It is known that the more general case (the so-called paramodular case) can be reduced to this strict Siegel case. As far as we know, the result seems new (and perhaps useful) even in the case $g=1$.

1. Stable subsets of $(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$. Let $m, n, \cdots$ denote elements of $\boldsymbol{R}^{2g}$ and $M, M', \cdots$ elements of $Sp(g,\boldsymbol{R})$. We shall denote by m' and m'' elements of $\boldsymbol{R}^g$ determined by the first and the last g coefficients of m; similarly for n' and n'' etc. Also we shall write M etc. in the form

$$M=\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \cdots$$

in terms of $g\times g$ submatrices $a, b, c, d, \cdots$. For a given pair (M, m), we define $M\cdot m$ by

$$M\cdot m=\begin{pmatrix} d & -c \\ -b & a \end{pmatrix} m+\tfrac{1}{2}\cdot\begin{bmatrix} (c\,{}^td)_0 \\ (a\,{}^tb)_0 \end{bmatrix},$$

in which $(c\,{}^td)_0$, for instance, means to take the vector determined by the diagonal coefficients of $c\,{}^td$. Observe that the coefficient matrix of m is ${}^tM^{-1}$

Received March 26, 1965.

[1] This work was partially supported by the National Science Foundation.

The content of this paper is a part of a lecture given at the Johns Hopkins University in the fall of 1964. Since that time, Mumford has greatly enriched this field. (Added in proof.)

and that both $c^t d$ and $a^t b$ are symmetric. Now, if M, M' have integer coefficients, i.e., if M, M' are contained in $Sp(g, \mathbf{Z}) = \Gamma_g(1)$, we have

$$MM' \cdot m \equiv M \cdot (M' \cdot m) \mod 1.$$

The verification is straightforward. Also we clearly have

$$\begin{aligned} &1_{2g} \cdot m = m \\ &m \equiv n \mod 1 \text{ implies } M \cdot m \equiv M \cdot n \mod 1. \end{aligned}$$

Therefore $\Gamma_g(1)$ operates on the set $(\mathbf{R}/\mathbf{Z})^{2g}$. Consequently, the stabilizer $\Gamma_g(1,2)$ of the neutral element $0 \mod 1$ of the additive group $(\mathbf{R}/\mathbf{Z})^{2g}$ is a subgroup of $\Gamma_g(1)$. In other words, the subset of $\Gamma_g(1)$ defined by the condition

$$(c^t d)_0 \equiv (a^t b)_0 \equiv 0 \mod 2$$

forms a subgroup. Since the matrix E in the Introduction is contained in $\Gamma_g(1,2)$, this subgroup is stable under $M \to {}^t M = E^{-1} M^{-1} E$. Therefore $\Gamma_g(1,2)$ can equally be defined by the condition

$$({}^t b d)_0 \equiv ({}^t a c)_0 \equiv 0 \mod 2.^2$$

We shall define a sequence of subgroups in which $\Gamma_g(1,2)$ is the first member.

Suppose that l is a positive integer. Then we define $\Gamma_g(l, 2l)$ as a subset of $\Gamma_g(1)$ by the following two conditions

$$\begin{aligned} &M \equiv 1_{2g} \mod l \\ &(c^t d)_0 \equiv (a^t b)_0 \equiv 0 \mod 2l. \end{aligned}$$

If we drop the second condition, we get the principal congruence group $\Gamma_g(l)$ of level l, and we have $\Gamma_g(l, 2l) = \Gamma_g(l) \cap \Gamma_g(1,2)$ for $l \equiv 1 \mod 2$. Therefore $\Gamma_g(l, 2l)$ for $l \equiv 1 \mod 2$ is a subgroup of $\Gamma_g(1)$ and so is $\Gamma_g(l, 2l)$ for $l \equiv 0 \mod 2$. In fact $\Gamma_g(l, 2l)$ for $l \equiv 0 \mod 2$ is a normal subgroup of $\Gamma_g(1)$ [cf. TC, Lemma 1]. Now, if m is in $\mathbf{Q}^{2g}$, clearly $M \cdot m$ is also in $\mathbf{Q}^{2g}$. Therefore $\Gamma_g(1)$ operates on the set $(\mathbf{Q}/\mathbf{Z})^{2g}$, and hence this set decomposes into "domains of transitivity" with respect to $\Gamma_g(1)$. In the following we shall determine them explicitly.

To begin with, we shall agree upon three standard notations. If G is an additive group and r a positive integer, we shall denote by ${}_rG$ the kernel of the "multiplication by r." In other words ${}_rG$ is the subgroup of G con-

---

[2] In Eichler [2], he defines $\Gamma_g(1,2)$ by this condition. He shows that it forms a multiplicative monoid and then, using a set of generators, he proves that it is a group. However, this second procedure seems to be redundant.

sisting of elements of G of order r. Also, if X, Y are subsets of G, we shall denote by $X \pm Y$ the subsets of G consisting of elements of the form $x \pm y$ with x in X and y in Y. Finally, we shall denote $\exp(2\pi i x)$ by $e(x)$.

Now $(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$ is an additive group. We shall denote by $\mathfrak{S}_r$ the subset of $(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$ consisting of elements of order exactly equal to r. We shall also denote by $\mathfrak{S}_\pm$ the subsets of ${}_2(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$ defined respectively by the condition $e(2{}^t m' m'') = \pm 1$. Since ${}_r(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$ decomposes into $\mathfrak{S}_d$'s for d dividing r, we see that the number of elements in $\mathfrak{S}_r$ is

$$r^{2g} \prod_{p \mid r} (1 - p^{-2g}),$$

in which p runs over prime factors of r. Also we know that the numbers of elements in $\mathfrak{S}_\pm$ are respectively given by

$$2^{g-1}(2^g \pm 1)$$

and both $\mathfrak{S}_+$ and $\mathfrak{S}_-$ are domains of transitivity with respect to $\Gamma_g(1)$. More generally, we have the following theorem:

THEOREM 1. *There exist three types of domains of transitivity in* $(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$ *with respect to* $\Gamma_g(1)$, *and they are*

$$\begin{array}{lll} \mathfrak{S}_r & r \equiv 0 & \mod 4 \\ \mathfrak{S}_r + \mathfrak{S}_\pm & r \equiv 1 & \mod 2. \end{array}$$

We shall use two lemmas both of which are well known:

LEMMA 1. *Let u be an element of $\boldsymbol{Z}^{2g}$ and r a positive integer such that the coefficients of u and r are relatively prime. Then there exists an element M of $\Gamma_g(1)$ such that its first column vector and u are congruent* $\bmod r$.

LEMMA 2. *Let M_0, M_1 be two elements of $\Gamma_g(1)$ and let r_0, r_1 be positive integers which are relatively prime. Then there exists an element M of $\Gamma_g(1)$ satisfying $M \equiv M_i \bmod r_i$ for $i = 0, 1$.*

We shall start proving the theorem. Let $\mathfrak{S}$ be an arbitrary domain of transitivity and let $m \bmod 1$ be a representative of $\mathfrak{S}$. We can assume that the order r of $m \bmod 1$ is even. In fact, if the order is odd, we take an element M of $\Gamma_g(1)$ not in $\Gamma_g(1,2)$ and replace $m \bmod 1$ by $M \cdot m \equiv {}^t M^{-1} m + M \cdot 0 \bmod 1$. We put $r = 2r_0$. Suppose first that r_0 is even. Then, for every M in $\Gamma_g(1)$, the order of ${}^t M^{-1} m \bmod 1$ is a multiple of 4 while the order of $M \cdot 0$ is at most 2. Hence ${}^t M^{-1} m$ and ${}^t M^{-1} m + M \cdot 0$ have the same order $\bmod 1$. Consequently $\mathfrak{S}$ is contained in $\mathfrak{S}_r$. We shall show that any two elements of Σ_r are equivalent. We have only to show that an arbitrary element $n \bmod 1$ of Σ_r is equivalent to $m \bmod 1$ with ${}^t m = (1/r\,0 \cdots 0)$. At any rate, by Lemma 1 we can assume that $n \equiv m \bmod 1/2$, i. e. that $2(n-m) \equiv 0 \bmod 1$. We shall settle the special case $g=1$. Put $n = m + 1/2 \cdot {}^t(q'q'')$ with q', q'' in $\boldsymbol{Z}$. Then

$$M = \begin{pmatrix} 1 + r_0 q' q'' & r_0 q'' \\ q' & 1 \end{pmatrix}$$

is an element of $\Gamma_1(1)$ and we have $M \cdot m \equiv n \bmod 1$. Next we shall consider the case $g \geqq 2$. Let n^* be the element of $\boldsymbol{Q}^2$ defined by ${}^t n^* = (n_1' n_1'')$ and let n^{**} be the element of $\boldsymbol{Q}^{2g-2}$ obtained from the remaining coefficients of n. If we define m^* and m^{**} similarly, we have ${}^t m^* = (1/r\,0)$ and $m^{**} = 0$. By the special case we have just treated, we can find an element M^* of $\Gamma_1(1)$ satisfying $M^* \cdot m^* \equiv n^* \bmod 1$. On the other hand $n^{**} \bmod 1$ is an element of ${}_2(\boldsymbol{Q}/\boldsymbol{Z})^{2g-2}$, hence it is either in Σ_+ or in Σ_- for $g-1$. If it is in Σ_+ for $g-1$, there exists an element M^{**} of $\Gamma_{g-1}(1)$ satisfying $M^{**} \cdot m^{**} \equiv n^{**} \bmod 1$. Therefore, if we define an element M of $\Gamma_g(1)$ by "composing" M^* and M^{**}, we have $M \cdot m \equiv n \bmod 1$. However, if $n^{**} \bmod 1$ is in Σ_- for $g-1$, there exists an element M^{**} of $\Gamma_{g-1}(1)$ such that

$$M^{**} \cdot {}^t(1/2\,0 \cdots 0\,1/2\,0 \cdots 0) \equiv n^{**} \bmod 1.$$

Therefore, if we define an element M' of $\Gamma_g(1)$ by composing M^* and M^{**}, we get $M' \cdot {}^t(1/r\,1/2 \cdots 0\,0\,1/2\,0 \cdots 0) \equiv n \bmod 1$. In this way, the problem is reduced to showing that the following two vectors

$$m = 1/r \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad n = m + 1/2 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

are equivalent mod 1 with respect to $\Gamma_2(1)$. However

$$M = \begin{bmatrix} 1 & -r_0 & r_0^2 & r_0 \\ 0 & 1 & r_0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & r_0 & 1 \end{bmatrix}$$

is an element of $\Gamma_2(1)$ and we have $M \cdot m \equiv n \bmod 1$. Therefore we have $\Sigma = \Sigma_r$. Suppose next that r_0 is odd. Then we can write $m \bmod 1$ uniquely in the form $m_0 + m_1 \bmod 1$ with $m_0 \bmod 1$ in Σ_{r_0} and $m_1 \bmod 1$ in Σ_2. Moreover, if M is in $\Gamma_g(1)$, we have $M \cdot m \equiv {}^t M^{-1} m_0 + M \cdot m_1 \bmod 1$. Therefore, if $m_1 \bmod 1$ is in Σ_+, i. e., if $m \bmod 1$ is in $\Sigma_{r_0} + \Sigma_+$, so is $M \cdot m \bmod 1$. Conversely, suppose that $n \equiv n_0 + n_1 \bmod 1$ is an arbitrary element of $\Sigma_{r_0} + \Sigma_+$. By Lemma 1 we can find an element M_0 of $\Gamma_g(1)$ satisfying ${}^tM_0^{-1}m_0 \equiv n_0 \bmod 1$. Also we can find an element M_1 of $\Gamma_g(1)$ satisfying $M_1 \cdot m_1 \equiv n_1 \bmod 1$. By Lemma 2 we can find an element M of $\Gamma_g(1)$ satisfying $M \equiv M_0 \bmod r_0$ and $M \equiv M_1 \bmod 2$. Then we have $M \cdot m \equiv n \bmod 1$. Therefore we have $\Sigma = \Sigma_{r_0} + \Sigma_+$. Clearly the same argument holds when Σ_+ is replaced by Σ_-. This completes the proof of Theorem 1.

COROLLARY. *A finite stable subgroup of* $(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$ *is of the form* ${}_r(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$ *with some even positive integer* r.

In fact, if Σ is a finite stable subset of $(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$, which is at the same time a subgroup, it contains the neutral element of $(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$, hence Σ_+. Since Σ_+ generates ${}_2(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$, therefore, this is contained in Σ. The rest follows from Theorem 1.

2. A formula of transformation of theta-functions. Let τ denote a point of the Siegel upper-half plane $\mathfrak{S}_g$ and let z be a point of $\boldsymbol{C}^g$. Then, for every m in $\boldsymbol{R}^{2g}$, the following series

$$\begin{aligned}\theta_m(\tau, z) = \sum_{p \in \boldsymbol{Z}^g} e[1/2 \cdot {}^t(p + m')\tau(p + m') \\ + {}^t(p + m')(z + m'')]\end{aligned}$$

defines a holomorphic function of the two variables τ and z, and it is called the *theta-function* of characteristic m. We shall not, however, denote this function by θ_m. In fact θ_m will be reserved for the *theta-constant* of characteristic m, i. e. the holomorphic function on $\mathfrak{S}_g$ defined by $\theta_m(\tau) = \theta_m(\tau, 0)$. We have

$$\theta_m(\tau, -z) = \theta_{-m}(\tau, z).$$

Moreover, if n is an element of $\boldsymbol{Z}^{2g}$, we have

$$\theta_{m+n}(\tau, z) = e({}^tm'n'')\theta_m(\tau, z).$$

These are obvious properties. On the other hand, if n is in $\boldsymbol{R}^{2g}$, a simple calculation shows that we have

$$\begin{aligned}\theta_m(\tau, z + (\tau 1_g)n) = e(-{}^tn'm'') \\ \cdot e(-1/2 \cdot {}^tn'\tau n' - {}^tn'(z + n''))\theta_{m+n}(\tau, z).\end{aligned}$$

In particular, if n is in $\boldsymbol{Z}^{2g}$, combining these two properties, we get

$$\begin{aligned}\theta_m(\tau, z + (\tau 1_g)n) = e({}^tm'n'' - {}^tm''n') \\ \cdot e(-1/2 \cdot {}^tn'\tau n' - {}^tn'z)\theta_m(\tau, z).\end{aligned}$$

Another special case is the case $m=0$. If we replace n by m in the so-specialized relation, we get

$$\theta_m(\tau, z) = e(1/2 \cdot {}^t m' \tau m' + {}^t m'(z + m'')) \\ \cdot \theta_0(\tau, z + (\tau 1_g) m).$$

On the other hand, we know that $Sp(g, \mathbf{R})$ operates on $\mathfrak{S}_g$ as

$$M \cdot \tau = (a\tau + b)(c\tau + d)^{-1}.$$

In fact $Sp(g, \mathbf{R})$ operates on the vector-bundle $\mathfrak{S}_g \times \mathbf{C}^g$ over $\mathfrak{S}_g$ as

$$M \cdot (\tau, z) = ((a\tau + b)(c\tau + d)^{-1}, {}^t(c\tau + d)^{-1} z).$$

If we take M from $Sp(g, \mathbf{Z}) = \Gamma_g(1)$, we have

$$\theta_{M \cdot 0}(M \cdot (\tau, z)) = \kappa(M) \det(c\tau + d)^{1/2} \\ \cdot e(1/2 \cdot {}^t z (c\tau + d)^{-1} c z) \theta_0(\tau, z),$$

in which $\kappa(M)$ is a constant. More precisely $\kappa(M)^2$ depends only on M and the sign of $\kappa(M)$ depends on the sign of $\det(c\tau + d)^{1/2}$. For the proof of this functional equation, we refer to TC, Section 2 or to a recent paper by Siegel [6]. If we replace z by $z + (\tau 1_g) m$ and use the relation we have recalled, we get a general functional equation

$$\theta_{M \cdot m}(M \cdot (\tau, z)) = \kappa(M) e(\phi_m(M)) \det(c\tau + d)^{1/2} \\ \cdot e(1/2 \cdot {}^t z (c\tau + d)^{-1} c z) \theta_m(\tau, z),$$

in which

$$\phi_m(M) = -1/2 \cdot ({}^t m' {}^t b d m' + {}^t m'' {}^t a c m'' - 2 {}^t m' {}^t b c m'' \\ - {}^t(a {}^t b)_0 (d m' - c m'')).$$

In particular, putting $z = 0$ we get

$$\theta_{M \cdot m}(M \cdot \tau) = \kappa(M) e(\phi_m(M)) \det(c\tau + d)^{1/2} \theta_m(\tau).$$

This functional equation is special. However, the functional equation of the theta-constants for $g+1$ implies the functional equation of theta-functions for g. In fact, if τ, z are given, we choose a point ω of $\mathfrak{S}_1$ with a large $\mathrm{Im}(\omega)$ so that

$$\tilde{\tau} = \begin{pmatrix} \tau & z \\ {}^t z & \omega \end{pmatrix}$$

will become a point of $\mathfrak{S}_{g+1}$. On the other hand, if M is given, we compose M with 1_2 to get an element $\tilde{M}$ of $\Gamma_{g+1}(1)$. Then the "components" τ', z', ω' of $\tilde{M} \cdot \tilde{\tau}$ are given by

$$(\tau', z') = M \cdot (\tau, z), \qquad \omega' = \omega - {}^t z (c\tau + d)^{-1} c z.$$

Moreover, if $\tilde{m}$ is composed of m and m_{g+1}', m_{g+1}'', then $\tilde{M}\cdot\tilde{m}$ is composed of $M\cdot m$ and m_{g+1}', m_{g+1}''. Also we have $\phi_{\tilde{m}}(\tilde{M}) = \phi_m(M)$. Therefore, if we take $m_{g+1}' = 0$ and compare the coefficients of $e(\omega/2)$ in the Fourier-Jacobi expansions introduced by Pyatetski-Shapiro [5] of $\theta_{\tilde{M}\cdot\tilde{m}}(\tilde{M}\cdot\tilde{\tau})$ and of $\theta_{\tilde{m}}(\tilde{\tau})$ in the functional equation for $\theta_{\tilde{M}\cdot\tilde{m}}(\tilde{M}\cdot\tilde{\tau})$, we get the functional equation for $\theta_{M\cdot m}(M\cdot(\tau, z))$.

Now we shall proceed to clarify two ambiguities concerning the functional equation or the *transformation-law*, one about the identical vanishing of θ_m and another about the constant $\kappa(M)$. The following theorem gives satisfactory information about the first point:

THEOREM 2. *Suppose that m is in $\boldsymbol{Q}^{2g}$. Then the theta-constant θ_m represents the constant 0 if and only if $m \bmod 1$ is in Σ_-.*

Proof. Since m is in $\boldsymbol{Q}^{2g}$, we can apply the uniqueness theorem of Fourier expansions to

$$\theta_m(\tau) = \sum_{p\in \boldsymbol{Z}^g} e({}^t(p+m')m'')\, e(1/2\cdot{}^t(p+m')\tau(p+m')).$$

In this way, we see that we have $\theta_m = 0$ if and only if $q + m' = -(p+m')$ has a solution q in $\boldsymbol{Z}^g$ for every p in $\boldsymbol{Z}^g$ and if

$$e({}^t(p+m')m'') + e(-{}^t(p+m')m'') = 0$$

holds. The condition simply means $2m' \equiv 0 \bmod 1$, $e(2\,{}^tm'm'') = -1$ and $e(2\,{}^tpm'') = 1$ for every p in $\boldsymbol{Z}^g$, namely $2m \equiv 0 \bmod 1$ and $e(2\,{}^tm'm'') = -1$. This completes the proof.

In the special case $g = 1$, we can say more. Namely, except for θ_m with $m \bmod 1$ in Σ_-, that means ${}^tm \equiv (1/2\ 1/2) \bmod 1$, none of the θ_m vanishes at any point of $\mathfrak{S}_1$. This follows from the fact that $\theta_0(\tau, z)$, in the case $g = 1$, has one and only one zero at $1/2\cdot(\tau+1)$ in $\boldsymbol{C} \bmod (\tau 1)\boldsymbol{Z}^2$.

We shall now obtain sufficient information about $\kappa(M)$. We need two lemmas:

LEMMA 3. *Both $\Gamma_g(1)$ and $\Gamma_g(2)$ are generated by matrices of the forms*

$$M = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}, \qquad {}^tM^{-1} = \begin{pmatrix} d & 0 \\ -b & a \end{pmatrix}.$$

Proof. The statement about generators of $\Gamma_g(1)$ follows from a well-known formulation (cf. TC, Theorem 6) and from

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

As for the statement about generators of $\Gamma_g(2)$, it is a weaker form of TC, Theorem 1. q. e. d.

LEMMA 4. *For every element* M *of* $\Gamma_g(1)$ *of the form*

$$M = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix},$$

we have

$$\kappa(M)^2 = \kappa({}^tM^{-1})^2 = \det(d).$$

In order to make the proof of this lemma as clear as possible, we recall the following well-known fact. Suppose that U is a complex manifold on which a group G operates as a group of holomorphic automorphisms. A non-vanishing holomorphic function R defined on U and depending on M in G is called a factor of automorphy if it satisfies $R(MM', \tau) = R(M, M' \cdot \tau) R(M', \tau)$ for all M, M' in G. If M is an arbitrary element of G and f a holomorphic function on U, clearly

$$(M \cdot f)(\tau) = R(M^{-1}, \tau)^{-1} f(M^{-1} \cdot \tau)$$

is a holomorphic function on U, and we have

$$M \cdot (M' \cdot f) = MM' \cdot f, \qquad \text{identity} \cdot f = f.$$

In this way G operates on the vector space of holomorphic functions on U. Now, if we take $U = \mathfrak{S}_g$, $G = Sp(g, \boldsymbol{R})$, then

$$R(M, \tau) = \det(c\tau + d)^k$$

defines a factor of automorphy for every integer k. We shall agree to use this factor of automorphy whenever we apply an element of $\Gamma_g(1)$ to a monomial of theta-constants of degree $2k$. Then, the transformation-law of theta-constants shows that we have

$$M^{-1} \cdot \prod_{\alpha=1}^{2k} \theta_{M \cdot m_\alpha} = \kappa(M)^{2k} e\Big(\sum_{\alpha=1}^{2k} \phi_{m_\alpha}(M) \Big) \prod_{\alpha=1}^{2k} \theta_{m_\alpha}$$

for every M in $\Gamma_g(1)$. In particular, if we take M from $\Gamma_g(1,2)$, we have

$$M^{-1} \cdot (\theta_0)^2 = \kappa(M)^2 (\theta_0)^2.$$

Since θ_0 is not the constant 0, therefore, we get $\kappa(MM')^2 = \kappa(M)^2 \kappa(M')^2$ for all M, M' in $\Gamma_g(1,2)$ and $\kappa(1_{2g})^2 = 1$. These being remarked, we shall start proving Lemma 4.

The fact that $\kappa(M)^2 = \det(d)$ follows from $\kappa(M)^2 \det(d) = 1$, which itself follows from the transformation-law $\theta_{M\cdot 0}(M\cdot\tau)^2=\kappa(M)^2\det(d)\theta_0(\tau)^2$ observing $\theta_{M\cdot 0}(M\cdot\tau)^2=\theta_0(\tau)^2$. In order to prove $\kappa(M)^2=\kappa({}^tM^{-1})^2$, we write M in the form

$$M=\begin{pmatrix}1 & s\\ 0 & 1\end{pmatrix}\begin{pmatrix}{}^td^{-1} & 0\\ 0 & d\end{pmatrix},$$

in which s is an integer symmetric matrix. The decomposition is unique and we have $a={}^td^{-1}$, $b=sd$. Observing ${}^tM=E^{-1}M^{-1}E$, we apply E, M^{-1}, E^{-1} in this order to $(\theta_0)^2$. Since E is in $\Gamma_g(1,2)$, we first get $E\cdot(\theta_0)^2=\kappa(E^{-1})^2(\theta_0)^2$. Since the solution of $M\cdot m=0$ is given by

$$m=-1/2\cdot\begin{bmatrix}0\\ {}^td(s)_0\end{bmatrix}$$

and since we have $\phi_m(M)=0$, we get $M^{-1}\cdot(\theta_0)^2=\kappa(M)^2(\theta_m)^2$. Finally, since the solution of $E\cdot n=m$ is given by

$$n=1/2\cdot\begin{bmatrix}{}^td(s)_0\\ 0\end{bmatrix}$$

and since we have $\phi_n(E)=0$, we get $E^{-1}\cdot(\theta_m)^2=\kappa(E)^2(\theta_n)^2$. Moreover, since E is in $\Gamma_g(1,2)$, we have $\kappa(E^{-1})^2\kappa(E)^2=\kappa(1_{2g})^2=1$. Therefore we get $({}^tM^{-1})^{-1}\cdot(\theta_0)^2=\kappa(M)^2(\theta_n)^2$. If we observe that we have ${}^tM^{-1}\cdot n=0$ and $\phi_n({}^tM^{-1})=0$, we get $({}^tM^{-1})^{-1}\cdot(\theta_0)^2=\kappa({}^tM^{-1})^2(\theta_n)^2$. Since θ_n is not the constant 0, we get $\kappa({}^tM^{-1})^2=\kappa(M)^2$. q. e. d.

THEOREM 3. *We have* $\kappa(M)^8=1$ *for all* M *in* $\Gamma_g(1)$ *and*

$$\kappa(M)^2=(-4/M)\ ^3$$

for all M *in* $\Gamma_g(2)$. *Moreover* κ^2 *is a character of* $\Gamma_g(1,2)$.

Proof. If $m \bmod 1$ is taken from Σ_+, we have

$$M^{-1}\cdot(\theta_{M\cdot m})^8=\kappa(M)^8(\theta_m)^8,$$

hence $\kappa(MM')^8(\theta_m)^8=\kappa(M)^8\kappa(M')^8(\theta_m)^8$. This implies $\kappa(MM')^8=\kappa(M)^8\kappa(M')^8$ for all M, M' in $\Gamma_g(1)$ and $\kappa(1_{2g})^8=1$. Therefore κ^8 is a character of $\Gamma_g(1)$. On the other hand, if $c=0$ in M, we have $\kappa(M)^4=\kappa({}^tM^{-1})^4=1$ by Lemma 4. According to Lemma 3, such M and ${}^tM^{-1}$ generate $\Gamma_g(1)$. Therefore we get $\kappa^8=1$ on $\Gamma_g(1)$. Moreover, the fact that κ^2 is a character of $\Gamma_g(1,2)$ has already been observed. In particular κ^2 is a character of $\Gamma_g(2)$. Since $(-4/M)$ also defines a character of $\Gamma_g(2)$, they will coincide on $\Gamma_g(2)$ if

---

[3] The definition of $(-4/M)$ in TC, p. 230 is not correct. For every M in $\Gamma_g(2)$, it is defined either as the product of $(-4/d_{ii})$ or as that of $(-4/a_{ii})$ for $i=1,2,\cdots,g$.

they coincide on a set of generators of $\Gamma_g(2)$. Suppose that M is an element of $\Gamma_g(2)$ in which $c=0$. Then we have $\kappa(M)^2=\det(d)$ by Lemma 4. On the other hand, since we have $d\equiv 1_g \bmod 2$, we get

$$\det(d)\equiv\prod_{i=1}^{g} d_{ii}\equiv\prod_{i=1}^{g}(-4/d_{ii}) \quad \bmod 4.$$

Therefore we have $\kappa(M)^2\equiv(-4/M) \bmod 4$, and, since both sides are 1 or -1, we get $\kappa(M)^2=(-4/M)$. We also have $\kappa({}^tM^{-1})^2=\kappa(M)^2=(-4/M)$ $=(-4/{}^tM^{-1})$ by Lemma 4. This completes the proof.

We note that the fact $\kappa^8=1$ on $\Gamma_g(1)$ is classical [cf. 1, 6]. Also the fact that κ^2 is a character of $\Gamma_g(1,2)$ is known [cf. 2]. However, as far as we are aware, the explicit formula $\kappa(M)^2=(-4/M)$ for M in $\Gamma_g(2)$ has been observed for the first time in TC. Since, however, we did not state it quite explicitly in TC and also since the proof we gave there is complicated, we have provided a simpler proof here. At any rate, as we shall see, this will give almost statisfactory information about monomials of theta-constants considered as modular forms.

3. The graded ring $\mathfrak{A}_g(r)$. Let Σ denote a finite stable subset of $(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$ with respect to $\Gamma_g(1)$. Then Σ is a disjoint union of a finite number of domains of transitivity we have discussed in Section 1. We shall consider the ring $\boldsymbol{C}[\Sigma]$ generated over $\boldsymbol{C}$ by θ_m with $m \bmod 1$ in Σ. Since the theta-constant θ_m is uniquely determined up to a root of unity by $m \bmod 1$, the ring is of finite type. If we assign $1/2$ to each θ_m as its weight, the ring is graded by the total degree. We shall denote by $\mathfrak{A}(\Sigma)$ the subring $\mathbf{C}[\Sigma]^{(2)}$, i.e. the graded ring generated over $\boldsymbol{C}$ by $\theta_m\theta_n$ with $m, n \bmod 1$ in Σ. Since Σ is stable with respect to $\Gamma_g(1)$, the transformation-law of theta-constants shows that $\Gamma_g(1)$ operates on $\mathfrak{A}(\Sigma)$ and, in this way, we get a homomorphism of $\Gamma_g(1)$ to the group of (degree-preserving) automorphisms of the graded ring $\mathfrak{A}(\Sigma)$. We shall denote by $\Gamma(\Sigma)$ the kernel of this homomorphism. An element M of $\Gamma_g(1)$, therefore, belongs to $\Gamma(\Sigma)$ if and only if it keeps $\theta_m\theta_n$ invariant for all $m, n \bmod 1$ in Σ. This condition can be expressed by $M^{-1}\cdot(\theta_{M\cdot m}\theta_{M\cdot n})=\theta_{M\cdot m}\theta_{M\cdot n}$, i.e. by

$$\theta_{M\cdot m}\theta_{M\cdot n}=\kappa(M)^2 e(\phi_m(M)+\phi_n(M))\theta_m\theta_n$$

for all $m, n \bmod 1$ in Σ. In particular, taking $n=m$ and extracting square roots, we get

$$\theta_{M\cdot m}=\kappa(M)e(\phi_m(M))\theta_m$$

for all $m \bmod 1$ in Σ. Since $\kappa(M)$ has a sign-ambiguity, we did not put $\pm$ in front of $\kappa(M)$. In the following, we shall determine $\Gamma(\Sigma)$ in the case Σ is also a subgroup of $(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$, i.e. in the case $\Sigma = {}_r(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$ for some even positive integer r.

We shall start by examining the Fourier expansion of θ_m more closely than we did in the proof of Theorem 2. We see immediately that there are three cases to be distinguished. The first is the case where $m' \equiv 0 \bmod 1$. Then we have

$$\theta_m(\tau) = 1 + \sum_p{}' (e({}^t pm'') + e(-{}^t pm''))e(1/2 \cdot {}^t p \tau p),$$

in which the summation is extended over representatives of $p \sim -p$ for $p \neq 0$ in $\boldsymbol{Z}^g$. Namely one and only one of p and $-p$ appears in the summation. The second is the case where $m' \not\equiv 0 \bmod 1$, but $2m' \equiv 0 \bmod 1$. Then we have

$$\theta_m(\tau) = e({}^t m'm'') \sum_p{}' \left(e({}^t pm'') + e(-{}^t(p+2m')m'')\right) \cdot e(1/2 \cdot {}^t(p+m')\tau(p+m')),$$

in which the summation is extended over representatives of $p \sim -p-2m'$ for all p in $\boldsymbol{Z}^g$. The third is the case where $2m' \not\equiv 0 \bmod 1$. Then we simply have

$$\theta_m(\tau) = e({}^t m'm'') \sum_{p \in \boldsymbol{Z}^g} e({}^t pm'') e(1/2 \cdot {}^t(p+m')\tau(p+m')).$$

Now, suppose that θ_n is another theta-constant which differs from θ_m by a non-zero constant factor. Then θ_n necessarily has the same type of Fourier expansions as θ_m. In fact, if both are the constant 0, they are of the second type. If neither θ_m nor θ_n is the constant 0, we have only to observe that the exponential functions $e(1/2 \cdot {}^t(p+m')\tau(p+m'))$ which appear in different types of expansions are never equal. These being remarked, we shall examine the condition under which M will be contained in $\Gamma(\Sigma)$ for $\Sigma = {}_r(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$. By what we have said, if we put $n = M \cdot m$, we have $n' \equiv \pm m' \bmod 1$, in which the sign may depend on m. At any rate, expressing n in terms of a, b, c, d and m', m'', we get

$$dm' - cm'' + 1/2 \cdot (c\,{}^t d)_0 \equiv \pm m' \quad \bmod 1.$$

This is true for all m satisfying $rm \equiv 0 \bmod 1$ except for those $m \bmod 1$ contained in Σ_-. If we take $m' = 0$, we can vary m'' freely and we get $c \equiv 0 \bmod r$. If we then take $m'' = 0$ and vary m' freely, we get $d \equiv \pm 1_g \bmod r$. Since $\Gamma(\Sigma)$ is a normal subgroup of $\Gamma_g(1)$, with M it also contains ${}^t M^{-1} = E^{-1}ME$.

Therefore, we get $b \equiv 0 \bmod r$ and $a \equiv \pm 1_g \bmod r$, hence $M \equiv \pm 1_{2g} \bmod r$. Moreover, if we put

$$M = \pm 1_{2g} + r \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

with $g \times g$ integer matrices $\alpha, \beta, \gamma, \delta$, the condition $\theta_{M \cdot m} = \kappa(M) e(\phi_m(M)) \theta_m$ implies

$$\kappa(M) e(1/2r \cdot {}^t x' {}^t \beta d x' - 1/2r \cdot {}^t x'' {}^t a \gamma x'' \mp 1/r \cdot {}^t x' \alpha x'') = 1$$

for $x' = rm'$ and $x'' = rm''$. The sign-ambiguity in the exponential factor corresponds to that in $M \equiv \pm 1_{2g} \bmod r$. This implies $\kappa(M) = 1$ and

$${}^t x' {}^t \beta d x' - {}^t x'' {}^t a \gamma x'' \mp 2 {}^t x' \alpha x'' \equiv 0 \quad \bmod 2r$$

for all x', x'' in $\boldsymbol{Z}^g$ not corresponding to $m \bmod 1$ in Σ_-. Since ${}^t\beta d$ is symmetric, by taking $x'' = 0$ and varying x' freely, we get ${}^t\beta d \equiv 0 \bmod r$ and $({}^t\beta d)_0 \equiv 0 \bmod 2r$. In the same way, we get ${}^t a \gamma \equiv 0 \bmod r$ and $({}^t a \gamma)_0 \equiv 0 \bmod 2r$. Furthermore, unless $r = 2$ we can take any pair of "unit vectors" in $\boldsymbol{Z}^g$ as (x', x''), and we get $\alpha \equiv 0 \bmod r$. However, in the case $r = 2$, we can only say that ${}^t x' \alpha x'' \equiv 0 \bmod 2$ whenever ${}^t x' x'' \equiv 0 \bmod 2$. By taking a pair of distinct unit vectors as (x', x''), we see that $\alpha \equiv$ diagonal matrix mod 2. Then, by taking a sum of two unit vectors as $x' = x''$, we get either $\alpha \equiv 0 \bmod 2$ or $\alpha \equiv 1_g \bmod 2$. We observe that, in the case $r = 2$, the two signs in the expression of M are both permissible. Therefore, if we choose the sign suitably, i.e. using the admissible change $\alpha \to \alpha \pm 1_g$, we get $\alpha \equiv 0 \bmod 2$. Since we have $\alpha + {}^t\delta \equiv 0 \bmod r$ in any case, we also get $\delta \equiv 0 \bmod r$. Consequently M satisfies the following two conditions

$$M \equiv \pm 1_{2g} \quad \bmod r^2$$
$$({}^t b d)_0 \equiv ({}^t a c)_0 \equiv 0 \quad \bmod 2r^2.$$

In particular, this shows that, if g is odd, the case $M \equiv -1_{2g} \bmod r^2$ has to be dropped. Otherwise we will get $\kappa(M)^2 = (-4/M) = (-1)^g = -1$, and this contradicts $\kappa(M) = 1$. On the other hand, if g is even, we see that -1_{2g} is in $\Gamma(\Sigma)$ for every Σ. It is also straightforward to see that $\Gamma_g(r^2, 2r^2)$ is contained in $\Gamma(\Sigma)$ for $\Sigma = {}_r(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$. Therefore, if we denote by $\mathfrak{A}_g(r)$ the graded ring $A(\Sigma)$ for $\Sigma = {}_r(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$, we have the following theorem:

THEOREM 4. *The normal subgroup of* $\Gamma_g(1)$ *consisting of elements which keep* $\mathfrak{A}_g(r)$ *element-wise invariant is*

$$\Gamma_g(r^2, 2r^2) \text{ or } \Gamma_g(r^2, 2r^2) \cup -\Gamma_g(r^2, 2r^2)$$

according as g is odd or even.

Now, if we denote by $A(\Gamma)$ the graded ring of Siegel modular forms belonging to a discrete subgroup Γ of $Sp(g, \boldsymbol{R})$, which is commensurable with $\Gamma_g(1)$, the theorem implies that $\mathfrak{A}_g(r)$ is a subring of $A(\Gamma)$ for $\Gamma = \Gamma_g(r^2, 2r^2)$. In general, if A is a graded integral domain, we shall denote by $F(A)$ the subfield of the field of fractions of A consisting of quotients of homogeneous elements of same degrees. We note that any (degree-preserving) automorphism of the graded ring A gives rise to an automorphism of $F(A)$. In particular, if we take $\mathfrak{A}(\mathfrak{S})$ as A, we get a homomorphism of $\Gamma_g(1)$ to the group of automorphisms of $F(\mathfrak{A}(\mathfrak{S}))$. Observe that the operation of $\Gamma_g(1)$ on $F(\mathfrak{A}(\mathfrak{S}))$ is given simply by $(M \cdot f)(\tau) = f(M^{-1} \cdot \tau)$. This being remarked, we shall compare $F(\mathfrak{A}_g(r))$ with $F(A(\Gamma))$ for $\Gamma = \Gamma_g(r^2, 2r^2)$. For this purpose, we shall determine the subgroup of $\Gamma_g(1)$ consisting of elements which keep $F(\mathfrak{A}_g(r))$ element-wise invariant. Since -1_{2g} certainly has this property, we see that $\Gamma_g(r^2, 2r^2) \cup -\Gamma_g(r^2, 2r^2)$ is contained in this group. Now suppose that M is an arbitrary element of the group. Then we have

$$\theta_{M \cdot m} \theta_n = e(\phi_m(M) - \phi_n(M)) \theta_m \theta_{M \cdot n}$$

for all m, n satisfying $rm \equiv rn \equiv 0 \bmod 1$. In particular, we can take $m, n \bmod 1$ from $\mathfrak{S}_+$. We apply the following element

$$N = \begin{pmatrix} 1_g & 2s \\ 0 & 1_g \end{pmatrix}$$

of $\Gamma_g(2)$ to the so-specialized relation, and we get

$$\begin{aligned} &{}^t(M \cdot m)' s (M \cdot m)' - {}^t m' s m' \\ &\qquad \equiv {}^t(M \cdot n)' s (M \cdot n)' - {}^t n' s n' \quad \bmod 1. \end{aligned}$$

Since s is an arbitrary $g \times g$ symmetric integer matrix, we can vary s as an integer diagonal matrix. Then, we get $2(M \cdot m)' + 2m' \equiv 2(M \cdot n)' + 2n' \bmod 2$, hence $({}^t M^{-1}(m+n))' \equiv (m+n)' \bmod 1$. In the same way, using ${}^t N^{-1}$ instead of N, we get $({}^t M^{-1}(m+n))'' \equiv (m+n)'' \bmod 1$, hence

$${}^t M^{-1}(m+n) \equiv m+n \quad \bmod 1$$

for all $m, n \bmod 1$ in $\mathfrak{S}_+$. This implies $M \equiv 1_{2g} \bmod 2$, hence $\theta_{M \cdot 0} = \theta_0$. If we put $n = 0$ in the original relation, therefore, we get $\theta_{M \cdot m} = e(\phi_m(M)) \theta_m$ for all m satisfying $rm \equiv 0 \bmod 1$. As we have seen in the course of the proof of Theorem 4, this implies that M is an element of $\Gamma_g(r^2, 2r^2) \cup -\Gamma_g(r^2, 2r^2)$. We have thus obtained the following result:

LEMMA 5. *The normal subgroup of* $\Gamma_g(1)$ *consisting of elements which keep* $F(\mathfrak{A}_g(r))$ *element-wise invariant is* $\Gamma_g(r^2, 2r^2) \cup -\Gamma_g(r^2, 2r^2)$.

We note that Baily observed in [1] that the subgroup in question lies somewhere between $\Gamma_g(r) \cup -\Gamma_g(r)$ and $\Gamma_g(2r^2) \cup -\Gamma_g(2r^2)$. We have now come to the point of revealing the real significance of theta-functions.

Suppose that τ is a point of $\mathfrak{S}_g$, z a point of $\mathbf{C}^g$ and k a non-negative integer. Then a holomorphic function $\Theta(z)$ of z satisfying the functional equation

$$\Theta(z+(\tau 1_g)n) = e(-k(1/2\cdot {}^t n'\tau n' + {}^t n' z))\Theta(z)$$

for every n in $\mathbf{Z}^{2g}$ is called a theta-function of order k. The set of such functions forms a complex vector space of dimension k^g. If r is a positive integer, clearly $\theta_m(\tau, rz)$ for $rm \equiv 0 \bmod 1$ is a theta-function of order $k = r^2$. Moreover, if we replace z by $z+(\tau 1_g)n$ with $rn \equiv 0 \bmod 1$, except for one same factor, it is multiplied by $e(r({}^t m' n'' - {}^t m'' n'))$. Since we obtain all characters of the additive group ${}_r(\mathbf{Q}/\mathbf{Z})^{2g}$ of $n \bmod 1$ in this way, the r^{2g} theta-functions $\Theta_m(z) = \theta_m(\tau, rz)$ are linearly independent over $\mathbf{C}$. Therefore, they form a base of the complex vector space of theta-functions of order r^2. We know, on the other hand, that the vector space of theta-functions of order k defines a closed imersion of the complex torus $\mathbf{C}^g/(\tau 1_g)\mathbf{Z}^{2g}$ to the complex projective space $P_N(\mathbf{C})$ with $N = k^g - 1$ for $k \geqq 3$. This is the so-called Lefschetz theorem and the reader finds its proof in Weil [7]. In particular, we have the following lemma:[4]

LEMMA 6. *Suppose that* τ *is any given point of* $\mathfrak{S}_g$. *Then at least one* θ_m *with* $rm \equiv 0 \bmod 1$ *does not vanish at* τ *for any given positive integer* $r \geqq 2$.

Of course, if we just want to prove this lemma, it is enough to know that the said vector space of theta-functions does define a morphism to the projective space, i. e. that not all theta-functions of order k vanish at any given point of $\mathbf{C}$, and the proof of this statement is straightforward and classical.

On the other hand, in the case $r = 2$, we have examined the defining ideal of the abelian variety in the projective space which is obtained by the immersion. In this way, we have established in TC, Section 5 the following lemma:

LEMMA 7. *The field* $F(\mathfrak{A}_g(2))$ *contains* $F(\Gamma_g(1))$.

We are now ready to prove the following generalized "fundamental lemma":

---

[4] This proof was suggested to us by Mr. Shioda, who also helped us in carefully reading the manuscript.

THEOREM 5. *The integral closure of the graded ring $\mathfrak{A}_g(r)$, which is generated over* **C** *by $\theta_m\theta_n$ with $rm \equiv rn \equiv 0 \bmod 1$ for any even positive integer r, in its field of fractions is $A(\Gamma_g(r^2, 2r^2))$, which is the graded ring of all Siegel modular forms belonging to $\Gamma_g(r^2, 2r^2)$.*

Proof. Theorem 4 shows that $\mathfrak{A}_g(r)$ is a subring of $A(\Gamma_g(r^2, 2r^2))$. Lemma 7 shows that $F(\mathfrak{A}_g(r))$ contains $F(\Gamma_g(1))$. Therefore $F(\mathfrak{A}_g(r))$ lies between $F(\Gamma_g(1))$ and $F(\Gamma_g(r^2, 2r^2))$. If we apply Galois theory to this situation using Lemma 5, we get $F(\mathfrak{A}_g(r)) = F(\Gamma_g(r^2, 2r^2))$. Since the graded ring $\mathfrak{A}_g(r)$ contains non-zero elements of degree 1, therefore, the fields of fractions of $\mathfrak{A}_g(r)$ and $A(\Gamma_g(r^2, 2r^2))$ coincide. Finally Lemma 6 can be used, as in the special case $r = 2$ treated in TC, to conclude that $A(\Gamma_g(r^2, 2r^2))$ is integral over $\mathfrak{A}_g(r)$. This completes the proof.

This theorem enables us to investigate the "unknown" ring $A(\Gamma_g(r^2, 2r^2))$ in terms of $\mathfrak{A}_g(r)$. We note that information concerning $A(\Gamma_g(r^2, 2r^2))$ will give information concerning $A(\Gamma_g(l))$ for any l because of Lemma 1 in TC and by the "going-down process" we explained in TC, Section 6 and exemplified in [4]. Here we shall state only the following immediate consequence of Theorem 5 (and of Theorem 3):

COROLLARY. *The graded ring $A(\Gamma_g(l))$ for any even positive integer l is the integral closure of the graded ring generated over* **C** *by monomials*

$$\theta_{m_1}\theta_{m_2}\cdots\theta_{m_{2k}}$$

of a certain type. More precisely, if we put

$$lm_\alpha = \begin{pmatrix} x_\alpha' \\ x_\alpha'' \end{pmatrix} \quad \alpha = 1, 2, \cdots, 2k,$$

the conditions are that the coefficients of x_α' and x_α'' are integers satisfying

$$\sum_\alpha (x_{\alpha i}')^2 \equiv \sum_\alpha (x_{\alpha i}'')^2 \equiv 0 \qquad \bmod 2l$$

$$\sum_\alpha x_{\alpha i}' x_{\alpha j}' \equiv \sum_\alpha x_{\alpha i}'' x_{\alpha j}'' \equiv 0 \qquad \bmod l$$

$$\sum_\alpha x_{\alpha i}' x_{\alpha j}'' \equiv 0 \quad (i \neq j) \qquad \bmod l$$

$$\sum_\alpha x_{\alpha i}' x_{\alpha i}'' \equiv \begin{cases} 0 & (l \equiv 0 \bmod 4) \\ lk/2 & (l \equiv 2 \bmod 4) \end{cases} \qquad \bmod l.$$

Now, it is true that $\mathfrak{A}_g(r)$ is not a "known" ring. However, we can at least obtain various information about this ring because it is constructed *explicitly*. In fact, we can say that *the problem is now of an algebraic nature.* Concerning this point, we propose the following two problems:

PROBLEM 1. *Investigate whether the graded ring* $\boldsymbol{C}[\Sigma]$ *for* $\Sigma = {}_r(\boldsymbol{Q}/\boldsymbol{Z})^{2g}$ *is integrally closed or not.*

PROBLEM 2. *Determine the ideal of relations of the ring-generators*

$$\theta_m \quad rm \equiv 0 \mod 1$$

of $\boldsymbol{C}[\Sigma]$, e.g. *exhibit a base of this ideal.*

We note that these two problems are, at least formally, similar to the first and the second main theorems in the classical invariant theory. Answers to these problems are known only in the cases $g = 1, 2$ for $r = 2$ [cf. 3, p. 243; 4, Section 1]. In both cases, the graded ring $\boldsymbol{C}[\Sigma]$ is integrally closed and "Riemann's theta-formulas" generate the ideal of relations of the ring generators θ_m with $m \bmod 1$ in Σ_+ of $\boldsymbol{C}[\Sigma]$.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] W. L. Baily, "On the theory of θ-functions, the moduli of abelian varieties, and the moduli of curves," *Annals of Mathematics*, vol. 75 (1962), pp. 342-381.

[2] M. Eichler, *Einführung in die Theorie der algebraischen Zahlen und Funktionen*, Birkhäuser Verlage Basel und Stuttgart (1963).

[3] J. Igusa, "On the graded ring of theta-constants," *American Journal of Mathematics*, vol. 86 (1964), pp. 219-246 (cited TC).

[4] ———, "On Siegel modular forms of genus two (II)," *American Journal of Mathematics*, vol. 86 (1964), pp. 392-412.

[5] I. I. Pyatetski-Shapiro, *The geometry of classical domains and the theory of automorphic functions*, Moscow (1961).

[6] C. L. Siegel, "Moduln Abelscher Funktionen," *Göttingen Nachrichten*, Nr. 25 (1964), pp. 365-427.

[7] A. Weil, *Introduction à l'études des variétés kähleriennes*, Actualités Scientifiques et Industrielles, Paris (1958).

1-1 CONTINUOUS MAPPINGS ONTO E^n.

By Leslie C. Glaser.[1]

I. Although the original motivation for some of the results developed in this paper was derived from the author's interest in the monotone mapping problem [1, 5], the main results tend in another direction and deal with contractible non-collapsible $(n-1)$-complexes in S^n obtained from images of certain n-manifolds with boundary.

It is known that if X is locally compact, $f(X) = E^n$ is monotone and is homeomorphic with the decomposition space X_f generated by f, then f is compact[2] (p. 3 [6]). If X is a locally compact separable metric space and f is a monotone map taking X onto E^n, then there exists a 1-1 continuous map of the decomposition space X_f generated by f onto E^n [4]. Along this line, the following question is of interest. How strongly can we restrict a space X and still get a 1-1 continuous map h of X onto E^n and yet still not have a homeomorphism? In [3], Proizvolov claimed to show that this is not possible with only some very mild restrictions on X. If all his results were correct one would have the monotone mapping problem as an immediate corollary.

The main results of this paper are that for $n \geqq 3$ there exist n-manifolds M with non-empty boundary in E^n and a pseudo-isotopy F_t taking M into E^n so that $F_0 =$ identity, F_t is a homeomorphism for $0 \leqq t < 1$, and F_1 is a 1-1 continuous map taking M onto E^n. It is also shown that under appropriate continuity conditions, the one-point compactification of $F_1(\mathrm{Bd}\, M)$ is a compact AR subset of S^n. If the one-point compactification of $F_1(\mathrm{Bd}\, M)$ can be taken as a subcomplex C of S^n, then C is a contractible non-collapsible $(n-1)$-complex.

In [2], the existence of such maps is demonstrated by giving several such examples for $n = 3$. In fact, each F_1 gotten is locally piecewise linear and the one-point compactification of $F_1(\mathrm{Bd}\, M^3)$ gives rise to a contractible non-collapsible 2-complex. Many of the well-known contractible 2-complexes

---

Received January 25, 1965.

[1] This paper is the substance of the first half of the author's Ph. D. thesis which was prepared under the supervision of Professor R. H. Bing at the University of Wisconsin.

[2] The terminology is explained in Section II.

are shown to be obtainable in this manner and additional such examples are gotten by using these methods. A relatively simple contractible 2-complex is also given that cannot be obtained in this manner.

II. The standard terminology N^n, S^n and I^n will denote euclidean n-space, the n-sphere and the unit n-cell, respectively. An n-manifold with boundary is a separable metric space each of whose points has a neighborhood whose closure is homeomorphic with I^n. We will use $\operatorname{int} M$ and $\operatorname{Bd} M$ to denote the interior and boundary of M respectively.

A mapping f taking X onto Y is monotone if counter images of points on Y are continua (that is, for $y \in Y$, $f^{-1}(y)$ is a compact connected set). A mapping f taking X onto Y is compact if for each compact set $A \subset Y$, $f^{-1}(A)$ is compact.

A subset R of a space S is a retract of S provided that there is a continuous mapping $r: S \to R$ such that $r(x) = x$ for each point x in R. A compact set X is an absolute retract, denoted AR, if and only if for each homeomorphism of X onto a subset $h(X)$ of a compact space Y, $h(X)$ is a retract of Y. A subset D of a space X is a deformation retract of X if there is a retraction r of X onto D which is homotopic to the identity mapping of X onto itself under a homotopy that leaves D fixed.

If U is an open subset of a compact metric space X and f is a continuous map taking U into a metric space Y, then f will be said to be continuous in U at $X - U$ if for any sequence $\{x_i\} \subset U$ converging to a point $x \in X - U$ such that $\{f(x_i)\}$ converges to $y \in f(U)$, it is also true that for any other sequence $\{\hat{x}_i\} \subset U$ converging to x, $\{f(\hat{x}_i)\}$ also converges to y.

III. Before we go into any of the results for $n \geqq 3$, let us consider the related problems for $n = 1$ and 2. The following theorem is elementary.

THEOREM 1. *If X is a connected, locally connected, locally compact, separable metric space and h is a* 1-1 *continuous map of X onto $E^1(I^1, S^1)$, then $X = E^1 (= I^1, = [0,1)$ or $S^1)$.*

Although the following theorem is probably not of interest in itself (and is surprisingly difficult to prove), its proof does illustrate the techniques and also gives some motivation for the results for $n \geqq 3$.

THEOREM 2. *Suppose $M^2 \subset I^2$ is a 2-manifold with boundary such that $\operatorname{int} M = \operatorname{int} I^2$, then there exists a* 1-1 *continuous map f with f continuous in M^2 at $I^2 - M^2$ taking M^2 onto $E^2(I^2)$ only if $M^2 = \operatorname{int} I^2 (M^2 = I^2)$. Furthermore, there exists no M^2 and f (as above) taking M^2 onto S^2.*

Suppose there exists a 2-manifold M^2 and f satisfying the hypothesis of the theorem such that $f(M^2) = E^2$ but $M^2 \neq \operatorname{int} I^2$. (Note we are assuming M^2 is an open subset of I^2 such that $\operatorname{int} I^2 \subset M^2$). Then each component of $\operatorname{Bd} M^2$ is an open interval. Using the fact that f is 1-1 and continuous it follows that $f(\operatorname{Bd} M^2)$ is a closed 1-dimensional non-compact subset of E^2 such that $E^2 - f(\operatorname{Bd} M^2) = E^2$. That is, $f(\operatorname{Bd} M^2)$ is closed in E^2 by invariance of domain and $E^2 - f(\operatorname{Bd} M^2) = E^2$ since $f|_{\operatorname{int} I^2}$ is a homeomorphism. Also, if $f(\operatorname{Bd} M^2)$ were compact then there would exist a simple closed curve $J \subset \operatorname{int} I^2 = \operatorname{int} M^2$ such that $\overline{f(\operatorname{Bd} M^2)} \subset \overline{\operatorname{int} f(J)}$. But then $f(M^2)$ would be contained in the compact set $\overline{\operatorname{int} f(J)} + \overline{f(\operatorname{int} J)}$ and this contradicts the hypothesis that f is onto E^2.

Let $f(\operatorname{Bd} M^2)^*$ denote the one-point compactification of $f(\operatorname{Bd} M^2)$ in the one-point compactification of E^2. That is,

$$f(\operatorname{Bd} M^2)^* = f(\operatorname{Bd} M^2) + \{w\} \subset E^2 + \{w\} = S^2.$$

For convenience, let us also denote the map taking M^2 onto $S^2 - \{w\}$ by f. Now $f(\operatorname{Bd} M^2)^*$ is a compact connected 1-dimensional set such that $S^2 - f(\operatorname{Bd} M^2)^* = E^2$. $f(\operatorname{Bd} M^2)^*$ is 1-dimensional since the addition of one point to a separable metric space does not increase its dimension. To show that $f(\operatorname{Bd} M^2)^*$ is connected, let us suppose otherwise. Let C be a component of $f(\operatorname{Bd} M^2)^*$ not containing the compactification point $\{w\}$. Then there is a simple closed curve J in $S^2 - f(\operatorname{Bd} M^2)^*$ separating C from the component of $f(\operatorname{Bd} M^2)^*$ containing $\{w\}$. $f^{-1}(J)$ is a simple closed curve in $\operatorname{int} I^2$. However, $f^{-1}(J)$ does not separate points of $\operatorname{Bd} M^2$ in M^2 and hence J does not separate points of $f(\operatorname{Bd} M^2)^*$.

We now make use of the assumption that f is continuous in M^2 at $I^2 - M^2$. This allows us to extend the map f taking M^2 onto $S^2 - \{w\}$ to a map f^* taking I^2 onto S^2. That is, if $\{x_i\}$ is a sequence of points of M^2 converging to $x \in \operatorname{Bd} I^2 - M^2$, then if $\{f(x_i)\} \to y \in f(\operatorname{Bd} M^2)$, define $f^*(x) = y$. If $\{f(x_i)\}$ has no limit point in E^2, define $f^*(x) = \{w\}$. It then follows from the definition of f being continuous in M^2 at $I^2 - M^2$ that f^* is well-defined. That is, for any sequence $\{x_i\}$ in M^2 converging to $x \in I^2$, $\{f^*(x_i)\} \to f^*(x) \in S^2$. The continuity of f^* is elementary.

We note that $f^*|_{\operatorname{Bd} I^2}$ is a continuous map taking $\operatorname{Bd} I^2$ onto $f(\operatorname{Bd} M^2)^*$. Let Σ be a canonical square in $\operatorname{int} I^2$ concentric with $\operatorname{Bd} I^2$ and denote their common center by p. $f^*(I^2 - \operatorname{int} \Sigma) = f(M^2 - \operatorname{int} \Sigma) + \{w\}$ is a disk D in S^2 containing $f(\operatorname{Bd} M^2)^*$ in its interior, with boundary $f^*(\operatorname{Bd} \Sigma) = f(\operatorname{Bd} \Sigma)$. Let g be the map taking $I^2 - \operatorname{int} \Sigma$ onto $\operatorname{Bd} I^2$ by projecting from p. Then $f^* \circ g \circ f^{*-1} = F$ is a map taking D onto $f(\operatorname{Bd} M^2)^*$ such that $F|_{f(\operatorname{Bd} M^2)^*} = \text{identity}$. It follows that $f(\operatorname{Bd} M^2)^*$ is a retract of D. Therefore,

$f(\mathrm{Bd}\,M^2)^*$ is a locally connected 1-dimensional compact contractible subset of S^2 and hence a dendrite (p. 88, [4]). Let $r \neq \{w\}$ be an end point of $f(\mathrm{Bd}\,M^2)^*$. Now $f^{-1}(r) \in \mathrm{Bd}\,M^2$ and hence lies in an open interval in $\mathrm{Bd}\,M^2$. But then r lies in an interval in $f(\mathrm{Bd}\,M^2)^*$. This leads to a contradiction that r was an end point of $f(\mathrm{Bd}\,M^2)^*$. Therefore, $\mathrm{Bd}\,M^2 = \phi$ and $M^2 = \mathrm{int}\,I^2$. This completes the first part of the theorem.

Now suppose f is a 1-1 continuous map taking M^2 onto I^2 with f continuous in M^2 at $I^2 - M^2$. It follows from invariance of domain that $\mathrm{Bd}\,I^2 \subset f(\mathrm{Bd}\,M^2)$. $f^{-1}(\mathrm{Bd}\,I^2)$ is a closed subset of $\mathrm{Bd}\,M^2$. Suppose $\mathrm{Bd}\,M^2 - f^{-1}(\mathrm{Bd}\,I^2) \neq \phi$. Then there exists a 1-1 continuous map of $M^2 - f^{-1}(\mathrm{Bd}\,I^2)$ onto $\mathrm{int}\,I^2$ such that $\mathrm{Bd}(M^2 - f^{-1}(\mathrm{Bd}\,I^2)) \neq \phi$. But then this would contradict the above; therefore $\mathrm{Bd}\,M^2 - f^{-1}(\mathrm{Bd}\,M^2) = \phi$ and $f|_{\mathrm{Bd}\,M^2}$ is a 1-1 continuous map taking $\mathrm{Bd}\,M^2$ onto $\mathrm{Bd}\,I^2$. Since $\mathrm{Bd}\,M^2$ and $\mathrm{Bd}\,I^2$ are both 1-manifolds without boundary, it follows that $\mathrm{Bd}\,M^2 = \mathrm{Bd}\,I^2$ and that $M^2 = I^2$.

Finally, suppose we have an f and M^2 such that f takes M^2 onto S^2. Clearly, $M^2 \neq \mathrm{int}\,I^2$ or I^2 and $\mathrm{Bd}\,M^2$ is an open proper subset of $\mathrm{Bd}\,I^2$. $f(\mathrm{Bd}\,M^2)$ is a compact subset of S^2 since it is closed. Let p be a non-separating point of $f(\mathrm{Bd}\,M^2)$ and $g = f^{-1}(p)$. Then $f|_{M^2-\{q\}}$ is a map taking $M^2 - \{q\}$ onto E^2 satisfying the conditions of the first part of the theorem and hence $M^2 - \{q\} = \mathrm{int}\,I^2$. But then $M^2 = \mathrm{int}\,I^2 + \{q\}$ where $q \in \mathrm{Bd}\,I^2$ and this contradicts the assumption that M^2 was an open subset of I^2. Therefore, no such f and M^2 exists. This completes the proof of Theorem 2.

Question 1. Is Theorem 2 true without the assumption that f is continuous in M^2 at $I^2 - M^2$?

Question 2. If X is a connected, locally compact, locally connected separable metric space and if there exists a 1-1 continuous map h taking X onto $E^2(I^2$ or $S^2)$, then is h a homeomorphism?

IV. For $n \geqq 3$, the situation is entirely different.

LEMMA 1. *Suppose M is an n-manifold with non-empty boundary such that* $\mathrm{int}\,M = \mathrm{int}\,I^n$ *and* $n \geqq 3$. *If there is a* 1-1 *continuous map f taking M onto E^n, then* $f(\mathrm{Bd}\,M)$ *is a non-compact closed set in E^n such that* $E^n - f(\mathrm{Bd}\,M)$ *is dense and topologically E^n.*

$f(\mathrm{Bd}\,M)$ is closed in E^n, otherwise, there would exist a point $p \in \mathrm{int}\,M$ such that $f(p) \in \overline{f(\mathrm{Bd}\,M)}$. But then, there would exist a ball $B \subset \mathrm{int}\,M$ with p in its interior and $f(\mathrm{int}\,B)$ would be an open set containing $f(p)$ missing $f(\mathrm{Bd}\,M)$. Therefore, $f(\mathrm{Bd}\,M)$ is closed.

Suppose $f(\mathrm{Bd}\,M)$ is compact. Then $f(\mathrm{Bd}\,M)$ lies in the interior of an

n-sphere S in E^n and $f^{-1}(S)$ is an n-sphere in $\text{int}\, M$. Since $\text{Bd}\, M \subset \text{ext}\, f^{-1}(S)$ (in M), we have $\overline{f(\text{ext}\, f^{-1}(S))}$ is a connected set lying in $\overline{\text{int}\, S}$. But then $f(M) \subset \overline{\text{int}\, S} + \overline{f(\text{int}\, f^{-1}(S))}$ which is a compact set. This contradicts the hypothesis that $f(M) = E^n$.

For any $p \in f(\text{Bd}\, M)$, $f^{-1}(p)$ lies in the interior of a closed $(n-1)$-ball D in $\text{Bd}\, M$ and $f(D)$ would be an $(n-1)$-ball in E^n. Since $f(\text{Bd}\, M)$ can be expressed as the countable union $(n-1)$-balls in E^n, $E^n - f(\text{Bd}\, M)$ is dense and $f(\text{Bd}\, M)$ is $(n-1)$-dimensional.

Finally $E^n - f(\text{Bd}\, M) = f(\text{int}\, M) = E^n$, since $\text{int}\, M = E^n$ and $f(\text{int}\, M)$ is open in E^n.

THEOREM 3. *Let M be a n-manifold $(n \geqq 3)$ with non-empty boundary for which there exists* 1-1 *continuous map f taking M onto E^n. If the one-point compactification of $f(\text{Bd}\, M)$ (denoted $f(\text{Bd}\, M)^*$) is a subcomplex C under some triangulation T of S^n (in the one-point compactification of E^n), then $S^n - f(\text{Bd}\, M)^* = E^n$ and there exists a deformation retract of $S^2 - \{p\}$ onto $f(\text{Bd}\, M)^*$, where $p \in S^n - f(\text{Bd}\, M)^*$, and C is a contractible non-collapsible $(n-1)$-complex.*

It follows from Lemma 1 that $S^n - f(\text{Bd}\, M)^* = E^n$ and that $f(\text{Bd}\, M)^*$ is $(n-1)$-dimensional. The complex C ($|C| = f(\text{Bd}\, M)^*$) is non-collapsible since each point $p \in f(\text{Bd}\, M)$ lies in an open $(n-1)$-ball in $f(\text{Bd}\, M)$ and hence C has no free edges. Since $f(\text{Bd}\, M)^*$ is a subcomplex C under the triangulation T of S^n, the open simplicial neighborhood U of C under the second barycentric subdivision T'' of T can be mapped by a deformation retraction onto C. That is, there exists a map $H: U \times I \to U$ such that:

$$\begin{aligned} &H = \text{identity on } U \times \{0\} \\ &H(U, 1) = |C| \\ \text{and } &H = \text{identity on } C \times \{t\},\ 0 \leqq t \leqq 1. \end{aligned}$$

Since $S^n - |C| = E^n$, for some $p \in S^n - |C|$ there exists a tame $(n-1)$-sphere Σ such that $\Sigma \subset U - (|C| + \{p\})$ and Σ separates $|C|$ from p.

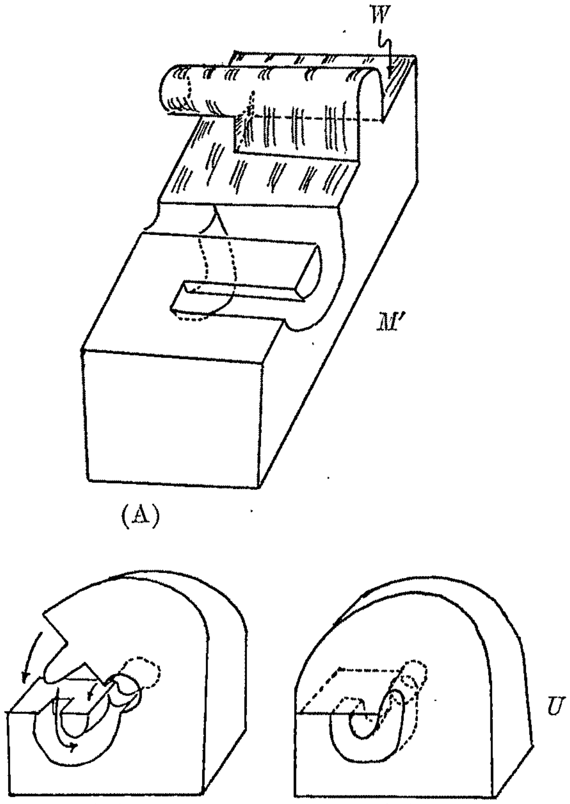
Let R be the component of $S^n - \Sigma$ containing $|C|$, then there exist a deformation retraction $K: (S^n - \{p\}) \times I \to S^n - \{p\}$ such that

$$\begin{aligned} &K = \text{identity on } (S^n - \{p\}) \times \{0\} \\ &K(S^n - \{p\}, 1) = \bar{R} \\ \text{and } &K = \text{identity on } \bar{R} \times \{t\},\ 0 \leqq t \leqq 1. \end{aligned}$$

Then $H \circ K$ is the desired deformation retraction of $S^n - \{p\}$ onto $|C|$. Therefore, C is a contractible non-collapsible $(n-1)$-complex of S^n.

THEOREM 4. *Let M be an n-manifold $(n \geqq 3)$ with non-empty boundary such that $M \subset I^n$ and $\text{int}\, M = \text{int}\, I^n$. Suppose f is a* 1-1 *continuous map* *taking M onto E^n such that f is continuous in M at $I^n - M$, then $f(\mathrm{Bd}\,M)^*$ is a $(n-1)$-dimensional compact AR subset of S^n.*

The proof of Theorem 4 follows exactly as in the first part of the proof of Theorem 2. That is, since f is continuous in M at $I^n - M$, f can be extended to a continuous map f^* taking I^n onto S^n, with $f^*(\mathrm{Bd}\,I^n) = f(\mathrm{Bd}\,M)^*$.



(A)

(B) (C)

FIGURE 1.

Also, we can get an n-cell N in S^n containing $f(\mathrm{Bd}\,M)^*$ in its interior and a retraction F taking N onto $f(\mathrm{Bd}\,M)^*$. Since $f(\mathrm{Bd}\,M)^*$ is a retract of N, it follows that $f(\mathrm{Bd}\,M)^*$ is a compact AR subset of S^n.

THEOREM 5. *For $n \geqq 3$, there exists n-manifolds M with non-empty boundary in E^n and a pseudo-isotopy F_t taking M into E^n so that $F_0 =$ identity, F_t is a homeomorphism for $0 \leqq t < 1$ and F_1 is a 1-1 continuous map taking M onto E^n.*

It will suffice to demonstrate the existence of such manifolds and pseudo-isotopies for $n=3$. That is, suppose M^3 is a 3-manifold with non-empty boundary in E^3 for which we have such a pseudo-isotopy F_t. For $n>3$, let $M^n = M^3 \times (\operatorname{int} I^{n-3})$. Then $\operatorname{int} M^n = \operatorname{int} M^3 \times \operatorname{int} I^{n-3} = \operatorname{int} I^n$ and $\hat{F}_t = F_t \times G_t$ will be a pseudo-isotopy satisfying the conclusion of the theorem, where G_t is an isotopy of $\operatorname{int} I^{n-3}$ into E^{n-3} with $G_0 = \mathrm{id}$ and $G_1(\operatorname{int} I^{n-3}) = E^{n-3}$. Here $\hat{F}_t(p) = \hat{F}_t(x,y) = (F_t(x), G_t(y))$, $p \in M^n$, $x \in M^3$, $y \in \operatorname{int} I^{n-3}$ and $p = (x,y)$.

In [2], the existence of such pseudo-isotopies for $n=3$ is given for a number of different examples and the resultant contractible 2-complex is discussed. For completeness, one of the simpler examples will be given here.

Let $M^3 = \operatorname{int} I^3$ plus an open annulus W in the top face of $\operatorname{Bd} I^3$. The pseudo-isotopy F_t of M^3 can be most easily described by indicating how we could deform M^3 in E^3 so as to obtain a set U which topologically E^3. If we continuously expand the deformed M^3 during the pseudo-isotopy we can get a pseudo-isotopy which gives us E^3 at the final stage.

By deforming M^3 as indicated in Figure 1, (A) to M' and then continuously folding M' as in Figure 1, (B) and (C), we get a set U homeomorphic to E^3. That is, M^3 is deformed to M' so that the open annulus W bounds a solid "tube" in M'. Note this tube is not closed at the left vertical end. In folding over M' we leave a tunnel in which the tube is inserted to get the resultant set U. Observe that the tunnel does not have a boundary while the tube does. Note that the one-point compactification of $F_1(\operatorname{Bd} M^3)$ is just the contractible non-collapsible 2-complex known as the dunce hat (that is a triangle abc with the three sides identified $ab = bc = ac$).

Rice University,
The University of Wisconsin.

REFERENCES.

[1] E. H. Connell, "Images of E^n under acyclic maps," *American Journal of Mathematics*, vol. 83 (1961), pp. 787-790.

[2] L. C. Glaser, *Contractible Complexes in S^n*, Ph. D. Thesis, University of Wisconsin, 1964.

[3] V. V. Proizvolov, "Contractions into euclidean spaces," *Soviet Mathematics Doklady*, vol. 4 (1963), pp. 1194-1195.

[4] G. T. Whyburn, *Analytic Topology*, American Mathematical Society Publication, 28 (1942).

[5] ———, "Compactness of certain mappings," *American Journal of Mathematics*, vol. 81 (1959), pp. 306-314.

[6] ———, "Decomposition spaces," *Topology of 3-manifolds*, Proceedings of the 1961 Topology Institute at the University of Georgia, Prentice-Hall, pp. 2-4.

DEFORMATIONS OF AFFINE EMBEDDINGS OF COMPLEX ANALYTIC HOMOGENEOUS SPACES.[1]

By G. Hochschild and G. D. Mostow.

---

1. **Introduction.** Let G be a complex analytic group whose points are separated by the complex analytic representative functions, and let L be a closed complex Lie subgroup of G. In [3], we have dealt with the existence of structures of quasi-affine algebraic varieties on the complex analytic homogeneous space $L\backslash G$ that are compatible with the canonical action of G on $L\backslash G$. Here we shall consider the problem of determining the totality of all such structures on a $L\backslash G$ when one such structure is given. We make the assumption that L has only a finite number of connected components, in which case the quasi-affine structures on $L\backslash G$ arise from the affine structures on G and are therefore under much better control. The main result is Theorem 3.2, which gives a direct construction yielding all quasi-affine structures from a given one. An important case is the case where L is trivial, in which case Theorem 3.2 describes how all the left stable basic subalgebras of the algebra of the representative functions are obtained from a given one. We also deal with the conjugacy of the quasi-affine structures with respect to the complex analytic automorphisms of G and, in particular, show that the conjugacy classes are in one to one correspondence with the orbits for the action of a group of linear automorphisms on a finite-dimensional complex vector space. The precise result is stated in Section 4, and the underlying parameterization of a representative set of quasi-affine structures is described in Proposition 3.1. Section 5 gives two concrete examples illustrating the construction of Theorem 3.2 and exhibiting extreme cases of the nature of the set of conjugacy classes of the affine structures of a complex analytic group.

2. **The toroidal setting.** Let G be a complex analytic group, and let $\boldsymbol{R}(G)$ denote the algebra of all complex analytic representative functions on G. *We assume throughout that* $\boldsymbol{R}(G)$ *separates the points of* G. The group of all proper automorphisms of $\boldsymbol{R}(G)$ is denoted by A, and G will be identified with its canonical image in A whenever convenient. If S is an A-stable

---

Received March 1, 1965.

[1] Research done with partial support of NSF Grants GP 1610 and GP 1662.

subalgebra of $\boldsymbol{R}(G)$ and B is a subgroup of A then B_S denotes the restriction image of B in the group of algebra automorphisms of S. If V is a subset of $\boldsymbol{R}(G)$ then A^V denotes the fixer of V in A. Similarly, with S and B as above, S^B denotes the B-fixed part of S. Finally, if f is an element of $\boldsymbol{R}(G)$ then f' denotes the element of $\boldsymbol{R}(G)$ given by $f'(x) = f(x^{-1})$ for every x in G, and S' denotes the image of S under the involution $f \to f'$. We say that a subalgebra of $\boldsymbol{R}(G)$ is *fully stable* if it is stable under the left and right actions of G and under the involution $f \to f'$.

Let L be a closed complex Lie subgroup of G such that L/L_1 is finite, where L_1 denotes the connected component of 1 in L, and suppose that $L \backslash G$ has a quasi-affine structure P in $\boldsymbol{R}(G)$, in the sense of [3]. The proof of Theorem 5.3 of [3] (saying that every quasi-affine structure on $L \backslash G$ is liftable to G) has shown that there is a fully stable finitely generated subalgebra S of $\boldsymbol{R}(G)$ satisfying the following requirements: (1) S contains an affine carrier for $L \backslash G$ in P; (2) S contains $\mathrm{Hom}(G, C)$ and separates the points of G; (3) there is a complex toroid $X(S)$ in $(A^{P'})_S$ such that A_S is the semidirect product $X(S) \cdot G_S$.

As is shown in the proof of Theorem 4.1 of [2], it follows from (2) that the kernel of the restriction epimorphism $A \to A_S$ lies in the center of A and consists entirely of semisimple automorphisms. Let X be the full inverse image of $X(S)$ in A, so that $X_S = X(S)$. Since S contains an affine carrier for $L \backslash G$ in P and X leaves the elements of S' fixed, we have $X \subset A^{P'}$. Clearly, $XG = A$. Since $X \cap G$ is trivial on S and S separates the points of G, we have $X \cap G = (1)$, so that A is the semidirect product $X \cdot G$. Let S_1 be any fully stable finitely generated subalgebra of $\boldsymbol{R}(G)$ containing S. Then X_{S_1} is evidently an abelian fully reducible algebraic subgroup of A_{S_1}. Moreover, A_{S_1} is the semidirect product $X_{S_1} \cdot G_{S_1}$, which shows that X_{S_1} is connected. Thus X_{S_1} is a complex toroid. Hence it is clear that X is the projective limit of the complex toroids X_{S_1}, which we shall express by saying that *X is a complex pro-toroid in A.*

The commutators $xyx^{-1}y^{-1}$, where x belongs to X and y to L, lie in $A^{P'} \cap G = L$. Hence X lies in the normalizer $N_A(L)$ of L in A. Clearly, $A^{P'}$ is the semidirect product $X \cdot L$, and, by Theorem 2.2 of [3], we have $P' = \boldsymbol{R}(G)^{A^{P'}} = \boldsymbol{R}(G)^{X \cdot L}$.

Choose a maximal complex toroid T_S in $N_{A_S}(L_S)$ containing X_S. Let T be the full inverse image of T_S in A, and let S_1 be any fully stable finitely generated subalgebra of $\boldsymbol{R}(G)$ containing S. Since the kernel of the restriction epimorphism $A_{S_1} \to A_S$ is contained in X_{S_1}, we have $T_{S_1}/X_{S_1} \approx T_S/X_S$, whence T_{S_1} is connected. Using the above information on the kernel of the map $A \to A_S$, one now sees easily that T is a maximal complex pro-toroid in

$N_A(L)$ containing X. Since A is the semidirect product $X \cdot G$, we have $T = X \times (T \cap N_G(L))$. Similarly, $T_S = X_S \times (T_S \cap N_G(L)_S)$. Since T is the full inverse image of T_S in A and the restriction $G \to G_S$ is an isomorphism, we have $T_S \cap N_G(L)_S = (T \cap N_G(L))_S$. Since T_S and X_S are complex toroids, so is therefore $T \cap N_G(L)$ (which may be identified with $(T \cap N_G(L))_S$, of course).

Let L^* denote the pro-algebraic hull of L in A, i.e., the projective limit of the system of algebraic hulls of the L_{S_1}'s for the fully stable finitely generated subalgebras S_1 of $\boldsymbol{R}(G)$. Clearly, $L^* \subset A^{P'} = X \cdot L$.

PROPOSITION 2.1. *Let G, L, L^*, X, T be as above. Let Y be any complex pro-toroid in A such that $T = Y \times (T \cap N_G(L))$ and $L^* \subset YL$. Then $(\boldsymbol{R}(G)^{YL})'$ is a quasi-affine structure for $L \backslash G$, and YL is the fixer of $\boldsymbol{R}(G)^{YL}$ in A.*

Proof. Let us write P_1 for $(\boldsymbol{R}(G)^{YL})'$. By Theorem 4.1 of [3], $\boldsymbol{R}(G)^Y$ is a basic subalgebra of $\boldsymbol{R}(G)$. In particular, we have therefore $\boldsymbol{R}(G)^Y = S_1{}^Y$, where S_1 is a finitely generated fully stable subalgebra of $\boldsymbol{R}(G)$. Now $P_1' = S_1{}^{YL}$, whence the field of fractions of P_1' is a subfield of the field of fractions of S_1, and is therefore finitely generated. Hence also the field of fractions of P_1 is finitely generated. Moreover, it is clear from the form of P_1 as the fixed part of $\boldsymbol{R}(G)$ with respect to a group of algebra automorphisms that the intersection of its field of fractions with $\boldsymbol{R}(G)$ coincides with P_1. Since $Y \subset A^{P_1'}$ and $YG = XG = A$, we have $A^{P_1'}G = A$. Thus the conditions of Theorem 2.3 of [3] are satisfied, whence we conclude that P_1 is a quasi-affine structure for $(A^{P_1'} \cap G) \backslash G$. What remains to be proved is that $A^{P_1'} \cap G = L$. Since $(YL) \cap G = L$, it suffices therefore to show that $A^{P_1'} = YL$, i.e., that the fixer of $\boldsymbol{R}(G)^{YL}$ in A is precisely YL.

Consider the fully stable finitely generated subalgebras V of $\boldsymbol{R}(G)$ containing the algebra S considered in the beginning. For each V, let $G(V)$ denote the subgroup of G consisting of all elements g such that g_V belongs to Y_V. The canonical map $G \to G_V$ is an isomorphism and sends $G(V)$ onto $Y_V \cap G_V$, which clearly coincides with $Y_V \cap N_G(L)_V$. Evidently, if $V_1 \subset V_2$ then $G(V_2) \subset G(V_1)$. Since $T_V \cap N_G(L)_V = (T \cap N_G(L))_V$, we have $Y_V \cap N_G(L)_V = Y_V \cap (T \cap N_G(L))_V$. Hence each $G(V)$ is an algebraic subgroup of the complex toroid $T \cap N_G(L)$. Now $\bigcap_V G(V) \subset Y \cap G = (1)$, because Y is a pro-algebraic subgroup of A. Since the $G(V)$'s are algebraic subgroups of the algebraic linear group $T \cap N_G(L)$, this evidently implies that $G(V) = (1)$, and so $Y_V \cap G_V = (1)$, for some V, and hence also for all larger V's. Hence we may choose V so that $Y_V \cap G_V = (1)$ and such that $(V^L)'$ separates the points of $L \backslash G$.

Let G' denote the commutator subgroup of G. Since $YL = YL^*$, it is clear that $(YL)_V$ is an algebraic subgroup of A_V. Hence $(YLG')_V$ is an algebraic subgroup of A_V, because it is generated by $(YL)_V$ and the algebraic group G'_V. The injection $G'_V \rightarrow (YLG')_V$ canonically induces an isomorphism of the algebraic variety $((YL)_V \cap G'_V)\backslash G'_V$ onto the algebraic variety $(YL)_V\backslash(YLG')_V$. Since $Y_V \cap G_V = (1)$, we have $(YL)_V \cap G'_V \subset L_V$, so that $(YL)_V \cap G'_V = L_V \cap G'_V = (L \cap G')_V$. Thus the algebraic variety $(YL)_V\backslash(YLG')_V$ is isomorphic with the algebraic variety $(L \cap G')_V\backslash G'_V$. Since the right L-fixed part of V separates the points of $L\backslash G$, the right $L \cap G'$—fixed part of the restriction of V to G' separates the points of $(L \cap G')_V\backslash G'_V$. By Theorem 4 of [1], this last algebraic variety is therefore quasi-affine, so that the isomorphic algebraic variety $(YL)_V\backslash(YLG')_V$ is also quasi-affine. Again by Theorem 4 of [1], this means that $(YL)_V$ is an *observable* algebraic subgroup of $(YLG')_V$, in the sense of [1]. Since $(YLG')_V$ is a normal algebraic subgroup of A_V, it is an observable subgroup of A_V, by Theorem 10 of [1]. The notion of observability has evidently the requisite transitivity property to conclude that $(YL)_V$ is an observable algebraic subgroup of A_V. By Theorem 4 of [1], it follows that $(V^{YL})'$ separates the points of $(YL)_V\backslash A_V$, so that the fixer of V^{YL} in A_V coincides with $(YL)_V$. Using that Y is the projective limit of the Y_V's and that $Y_V \cap G_V = (1)$ for each V as used above, one sees readily that YL is the projective limit of the $(YL)_V$'s. Hence we conclude from the last result that the fixer of $\boldsymbol{R}(G)^{YL}$ in A coincides with YL, which completes the proof of Proposition 2.1.

Let Z be any maximal complex pro-toroid in $N_A(L)$, and let V be a fully stable finitely generated subalgebra of $\boldsymbol{R}(G)$ such that A^V consists entirely of semisimple automorphisms in the center of A. Then one sees easily that Z_V is a maximal complex toroid in $N_A(L)_V$. Moreover, the connectedness argument we made before for T shows that Z is the full inverse image of Z_V in A. Since both Z_V and T_V are maximal complex toroids in the algebraic group $N_A(L)_V$, they are conjugate by an inner automorphism effected by an element of the commutator subgroup of $N_A(L)_V$, which lies in $N_G(L)_V$. Hence there is an element g in $N_G(L)$ (actually, in $N_G(L)'$) such that $T_V = (gZg^{-1})_V$. Since T and gZg^{-1} are maximal complex pro-toroids in $N_A(L)$, this implies, by what we have said above (about Z), that $T = gZg^{-1}$. *Thus any two maximal complex pro-toroids in $N_A(L)$ are conjugate by an inner automorphism effected by an element of $N_G(L)'$.*

The quasi-affine structures of $L\backslash G$ that are associated with Z in the same way in which the structure $P_1 = (\boldsymbol{R}(G)^{YL})'$ of Proposition 2.1 is associated with T are evidently the structures $P_1 \cdot g$. Thus *every quasi-affine structure* *for* $L\backslash G$ *is of the form* $P_1 \cdot g$, where P_1 *is associated with* T *and* g *is an element of* $N_G(L)$; *evidently, every such* $P_1 \cdot g$ *is a quasi-affine structure for* $L\backslash G$.

3. Parameterization. Let T denote the maximal complex pro-toroid in $N_A(L)$ that we used in Section 2, and let X be a complex pro-toroid such that $T = X \times (T \cap N_G(L))$, A is the semidirect product $X \cdot G$, and $(\boldsymbol{R}(G)^{XL})'$ is the given quasi-affine structure P for $L\backslash G$. In order to find all the quasi-affine structures P_1 that are associated with T, it suffices to find all the complex pro-toroids Y in $N_A(L)$ such that $L^* \subset YL$ and $T = Y \times (T \cap N_G(L))$. Given such a Y, we obtain a map τ of X into $T \cap N_G(L)$ that is determined by the condition that $x\tau(x)$ be the component of x in Y, with respect to the decomposition $T = Y \times (T \cap N_G(L))$, for every element x of X. Then τ is evidently a homomorphism. We know that, for every sufficiently large fully stable finitely generated subalgebra V of $\boldsymbol{R}(G)$, T_V is the direct product $Y_V \times (T \cap N_G(L))_V$, in the sense of algebraic linear groups. It follows that our homomorphism τ is *rational* in the sense that if f is any polynomial function on the algebraic group $T \cap N_G(L)$ then there is an element h of $\boldsymbol{R}(G)$ such that $f(\tau(x)) = x(h)(1)$ for every element x of X.

By Theorem 4.1 of [2], $\boldsymbol{R}(G)^X$ is a basic subalgebra of $\boldsymbol{R}(G)$, so that $\boldsymbol{R}(G) = \boldsymbol{R}(G)^X[Q]$, where $Q = \exp(\mathrm{Hom}(G, C))$. Hence it is clear that the above element h may be taken as a C-linear combination of elements of Q. If f is a rational *homomorphism* of $T \cap N_G(L)$ into C^* it follows from a familiar argument, using that the canonical map of $C[Q]$ into the algebra of all complex valued functions on X is injective, that h is then a uniquely determined element of Q. Denoting it by $\tau^*(f)$, we have $f(\tau(x)) = x(\tau^*(f))(1)$ for every x in X, and τ^* is evidently a homomorphism of the group $\mathrm{Hom}(T \cap N_G(L), C^*)$ (of all rational homomorphisms of $T \cap N_G(L)$ into the multiplicative group C^* of the non-zero complex numbers) into Q. Conversely, if α is any such homomorphism then, by the rational duality theorem for the complex toroid $T \cap N_G(L)$, there is, for every x in X, a uniquely determined element $\alpha^0(x)$ in $T \cap N_G(L)$ such that $f(\alpha^0(x)) = x(\alpha(f))(1)$ for every element f of $\mathrm{Hom}(T \cap N_G(L), C^*)$. One verifies directly that α^0 is a rational homomorphism (in the sense defined above) of X into $T \cap N_G(L)$, that $(\alpha^0)^* = \alpha$, and that $(\tau^*)^0 = \tau$.

For any rational homomorphism τ of X into $T \cap N_G(L)$, denote by X_τ the complex pro-toroid in T consisting of the elements $x\tau(x)$ with x in X. Clearly, we have then $T = X_\tau \times (T \cap N_G(L))$. Now suppose that $L^* \subset X_\tau L$. Let x be an element of $X \cap L^*$. We may write $x = x_1\tau(x_1)z$, with x_1 in X and z in L. But this gives $z \in T \cap L$, whence $x = x_1$ and $\tau(x) = z^{-1} \in L$.

Thus $\tau(X \cap L^*) \subset T \cap L$. Conversely, if τ satisfies this last condition we find that $L^* \subset X_\tau L$. For if t is any element of L^* we may write $t = xz$, with z in L and x in $X \cap L^*$. Now $x\tau(x) \in X_\tau$ and $\tau(x) \in L$, whence $t = xz \in X_\tau L$. Thus we have $L^* \subset X_\tau L$ if and only if $\tau(X \cap L^*) \subset T \cap L$.

We wish to translate this property of τ into the corresponding property of τ^*. It follows immediately from the definition of τ^* and the rational duality theorem for the complex toroid $T \cap N_G(L)$ that we have $\tau(X \cap L^*) \subset T \cap L$ if and only if τ^* sends the annihilator of $T \cap L$ in $\mathrm{Hom}(T \cap N_G(L), C^*)$ into $Q^{X \cap L^*}$. We shall now show that $Q^{X \cap L^*} = Q^L$, so that our condition on τ^* is actually independent of the choice of X. It is not difficult to see (in much the same way in which we proved the existence of T) that X is contained in a maximal reductive pro-algebraic subgroup M of A, where 'reductive' means that each M_V *is* reductive. Then we have $M = X(M \cap G)$, and $M \cap G$ is a maximal reductive subgroup, H say, of G. Also, $A = MU$, where U is the *unipotent hull* of G in A, i.e., the maximum normal subgroup of A that is unipotent on $\boldsymbol{R}(G)$. Thus $A = XHU$.

Let V be a fully stable finitely generated subalgebra of $\boldsymbol{R}(G)$ containing the subalgebra S used in Section 2. We know that $(A^{P'})_V$ is the semidirect product $X_V \cdot L_V$ and that L^*_V is contained in $(A^{P'})_V$. Hence L^*_V is the semidirect product $(X \cap L^*)_V \cdot L_V$. Clearly, $(X \cap L^*)_V$ is contained in a maximal fully reducible subgroup $M(L^*_V)$ of L^*_V, and we have a semidirect product decomposition $M(L^*_V) = (X \cap L^*)_V \cdot (M(L^*_V) \cap L_V)$. It follows from Lemma 5.2 of [3] that $M(L^*_V) \cap L_V$ is reductive. We may view this group as a closed complex Lie subgroup of G. Since it is reductive, it is contained in HN, where N is the radical of G'. Since $N \subset U$, we have therefore $M(L^*_V) \cap L_V \subset (HU)_V$, so that $M(L^*_V) \subset (X \cap L^*)_V(HU)_V$.

Now we have $L^*_V = M(L^*_V)U(L^*_V)$, where $U(L^*_V)$ is the maximum unipotent normal subgroup of L^*_V. Since $A = XHU$, the factor group $A_V/(HU)_V$ is a rational homomorphic image of X_V and hence has no unipotent elements other than 1. Hence we must have $U(L^*_V) \subset (HU)_V$. Together with our above result on $M(L^*_V)$, this gives $L^*_V \subset (X \cap L^*)_V(HU)_V$.

Since HU leaves the elements of Q fixed, we have therefore $Q^L = Q^{L^*} \subset Q^{X \cap L^*} \subset Q^{L^*}$, so that $Q^{X \cap L^*} = Q^L$. Thus we have $L^* \subset X_\tau L$ if and only if τ^* sends the annihilator of $T \cap L$ in $\mathrm{Hom}(T \cap N_G(L), C^*)$ into Q^L.

Since the quasi-affine structure corresponding to X_τ actually depends only on $X_\tau L$, we must now examine the condition $X_\tau L = X_\sigma L$, where σ and τ are rational homomorphisms of X into $T \cap N_G(L)$ sending $X \cap L^*$ into $X \cap L$. If $X_\tau L = X_\sigma L$ then, for every x in X, there is an element x_1 in X such that $x\sigma(x) = x_1\tau(x_1)z$, with z in L. As above, it follows from this that $x_1 = x$ and $\sigma(x)^{-1}\tau(x) \in L$. The last condition is equivalent to the condition that

$(\sigma^{-1}\tau)^*$ send the annihilator of $T \cap L$ in $\mathrm{Hom}(T \cap N_G(L), C^*)$ into $Q^X = (1)$, where σ^{-1} stands for the reciprocal of σ. Thus we have $X_\tau L = X_\sigma L$ if and only if σ^* coincides with τ^* on the annihilator of $T \cap L$, i.e., if and only if the restrictions of σ^* and τ^* to the subgroup $\mathrm{Hom}((T \cap N_G(L))/(T \cap L), C^*)$ of $\mathrm{Hom}(T \cap N_G(L), C^*)$ coincide. Since Q^L is divisible, every homomorphism of this subgroup into Q^L extends to a homomorphism of $\mathrm{Hom}(T \cap N_G(L), C^*)$ into Q^L. Hence we may state our result as follows.

PROPOSITION 3.1. *Let G be a complex analytic group such that $\mathbf{R}(G)$ separates the points of G, and let L be a closed complex Lie subgroup of G whose component group is finite. Suppose that $L\backslash G$ has a quasi-affine structure, and let T be a maximal complex pro-toroid in $N_A(L)$. Let $K(T,L)$ denote the complex toroid $(T \cap N_G(L))/(T \cap L)$. Then the map of* $\mathrm{Hom}(\mathrm{Hom}(K(T,L), C^*), Q^L)$ *that sends each element γ of this group onto $(\mathbf{R}(G)^{LX_{\gamma^0}})'$ is a bijection onto the set of all quasi-affine structures of $L\backslash G$ that are associated with T.*

Note that Q has the structure of a finite-dimensional complex vector space, isomorphic with $\mathrm{Hom}(G, C)$, and that Q^L is a vector subspace of Q. Since $K(T,L)$ is a complex toroid, $\mathrm{Hom}(K(T,L), C^*)$ is a finitely generated free abelian group. Hence *the bijection of Proposition* 3.1 *is a parameterization by the points of a finite-dimensional complex vector space.*

Given a quasi-affine structure P for $L\backslash G$ in $\mathbf{R}(G)$, our above results enable us to determine the set of all quasi-affine structures explicitly in terms of P and a maximal complex pro-toroid T in $N_A(L)$ with which P is associated. Let X be as before, so that $P = (\mathbf{R}(G)^{XL})'$. For every homomorphism $\gamma\colon \mathrm{Hom}(K(T,L), C^*) \to Q^L$, let γ^0 denote the corresponding rational homomorphism of X into $T \cap N_G(L)$, and let $\gamma(P) = (\mathbf{R}(G)^{LX_{\gamma^0}})'$ be the corresponding quasi-affine structure of $L\backslash G$. We give an explicit construction yielding $\gamma(P)$ from P and γ.

Since $T \cap N_G(L)$ normalizes XL, we may regard $P' = \mathbf{R}(G)^{XL}$ as a $T \cap N_G(L)$-module, with respect to the action of $T \cap N_G(L)$ from the left. As such, let us decompose P' into a direct sum $\sum_\rho P'_\rho$, where P'_ρ is the submodule corresponding to the rational complex character ρ of the complex toroid $T \cap N_G(L)$, so that $t \cdot f = \rho(t) f$ for every f in P'_ρ and every t in $T \cap N_G(L)$. Since P'_ρ lies in $\mathbf{R}(G)^L$, every ρ for which $P'_\rho \neq (0)$ must be trivial on $T \cap L$, so that $\gamma(\rho)$ is defined. Put $P^\gamma = \sum_\rho \gamma(\rho)^{-1} P'_\rho$. We claim that $(P^\gamma)' = \gamma(P)$, i.e., that $P^\gamma = \mathbf{R}(G)^{LX_{\gamma^0}}$.

Since $\gamma(\rho)^{-1} \in Q^L$, it is clear that $P^\gamma \subset \mathbf{R}(G)^L$. Now let x be an element of X, and consider the action of $x\gamma^0(x)$ on P^γ. If f is any element of P'_ρ we have

$$\begin{aligned} x\gamma^0(x)(\gamma(\rho)^{-1}f) &= x[\gamma^0(x)\cdot\gamma(\rho)^{-1}]x[\gamma^0(x)\cdot f] \\ &= x[\gamma(\rho)(\gamma^0(x))^{-1}\gamma(\rho)^{-1}]x[\rho(\gamma^0(x))f]. \end{aligned}$$

Now $\gamma^0(x)\in T\cap N_G(L)$, which is a reductive subgroup of G. Hence $q(\gamma^0(x))=1$ for every element of Q, so that the above is equal to $x(\gamma(\rho))^{-1}\rho(\gamma^0(x))x(f)$. But $x(f)=f$ and $\rho(\gamma^0(x))=x(\gamma(\rho))(1)$. Hence the last expression reduces to $\gamma(\rho)^{-1}f$. Thus we have shown that $P^\gamma\subset\boldsymbol{R}(G)^{XL_{\gamma^0}}$. If we interchange the roles of P and $\gamma(P)$ in the above we find that $\gamma(P)^{\gamma^{-1}}\subset\boldsymbol{R}(G)^{LX}=P'$, whence $\gamma(P)'\subset P^\gamma$. Thus $\gamma(P)'=P^\gamma$, or $\gamma(P)=(P^\gamma)'$. We may summarize our results as follows.

THEOREM 3.2. *Let G, L and T be as in Proposition 3.1, and let P be a quasi-affine structure of $L\backslash G$ that is associated with T. Let $P'=\sum_\rho P'_\rho$ be the decomposition of the $T\cap N_G(L)$-module P' into the direct sum of the characteristic submodules P'_ρ corresponding to the rational complex characters ρ of $K(T,L)=(T\cap N_G(L))/(T\cap L)$. For every homomorphism γ of* Hom$(K(T,L),C^*)$ *into Q^L, put $\gamma(P)=\sum_p\gamma(\rho)(P'_\rho)'$. Then the map $\gamma\to\gamma(P)$ is a bijection of* Hom(Hom$(K(T,L),C^*),Q^L)$ *onto the set of quasi-affine structures of $L\backslash G$ associated with T, and the translates $\gamma(P)\cdot g$, with g ranging over $N_G(L)'$, make up the set of all quasi-affine structures of $L\backslash G$.*

It is easy to see that the map of P into $\gamma(P)$ that sends each element f of $(P'_\rho)'$ onto $\gamma(\rho)f$ is an algebra isomorphism of P onto $\gamma(P)$. Hence it is clear that, *as C-algebras, the action of G being disregarded, all quasi-affine structures for $L\backslash G$ are isomorphic.*

4. Conjugacy under complex analytic automorphisms. To begin with, there are some immediate consequences of Theorem 3.2 that are worth mentioning. If either $K(T,L)$ or Q^L is trivial then Theorem 3.2 shows immediately that all quasi-affine structures of $L\backslash G$ are conjugate under an inner automorphism of G effected by an element of $N_G(L)'$. Clearly, $K(T,L)$ is trivial whenever the connected component of 1 in $N_G(L)$ lies in L, and also whenever G is simply connected and solvable, in which case there can be no non-trivial complex toroid in G.

On the other hand, Q^L is trivial if and only if $G=LHN$, where H is a maximal reductive subgroup of G and N is the radical of G'. In order to see this, let F denote the fixer of Q^L in G. Then F is also the fixer of Hom$(G,C)^L$ in G. It is known that the fixer of Hom(G,C) in G is precisely HN and that G/HN is a vector group. We may identify Hom(G,C) with Hom$(G/HN,C)$. Now $HN\subset F$, and F/HN is the fixer in G/HN of

$\mathrm{Hom}(G/HN, C)^{(LHN)/HN}$. Since L/L_1 is finite, it is clear that $(LHN)/HN$ coincides with the vector subgroup $(L_1HN)/HN$ of G/HN. Hence $F/HN = (LHN)/HN$, so that $F = LHN$. Evidently, this establishes our above claim.

We shall now discuss the conjugacy question in general. Let W denote the group of all complex analytic automorphisms of G. An element η of W extends canonically to an automorphism of A such that, for α in A and f in $\boldsymbol{R}(G)$, $\eta(\alpha)(f) = \alpha(f \circ \eta) \circ \eta^{-1}$. In this way, we shall regard W also as a group of rational automorphisms of the pro-algebraic group A. Let $W(T, L)$ denote the intersection of the stabilizers of T and L in W. Clearly, $W(T, L)$ acts as a group of permutations on the set of the subgroups $X_\tau L$. The corresponding action on the quasi-affine structures is the natural composition action; if $\eta(X_\tau L) = X_\sigma L$ then $\sigma^*(P) = \tau^*(P) \circ \eta^{-1}$. Writing $\eta^*(f)$ for $f \circ \eta^{-1}$, the corresponding action of $W(T, L)$ on $\mathrm{Hom}(\mathrm{Hom}(K(T, L), C^*), Q^L)$ is given by $\gamma \to (\eta^*) \circ \gamma \circ (\eta^{-1})^*$. In particular, $W(T, L)$ thus acts linearly with respect to the vector space structure of $\mathrm{Hom}(\mathrm{Hom}(K(T, L), C^*), Q^L)$.

A priori, it seems possible that two quasi-affine structures associated with T are conjugate under W without being conjugate under the smaller group $W(T, L)$. We shall show, however, that this is not the case. Let η be an element of W which transforms some structure $(\boldsymbol{R}(G)^{X_\tau L})'$ associated with T onto another such structure. Since $(X_\tau L) \cap G = L = \eta(X_\tau L) \cap G$, it is clear that η must lie in the stabilizer of L. Let $Z(X_\tau)$ denote the group of all elements z in $N_A(L)$ such that $zxz^{-1}x^{-1} \in L$ for all elements x of X_τ. Then we have evidently $\eta(Z(X_\tau)) = Z(\eta(X_\tau))$. Hence both T and $\eta(T)$ are maximal complex pro-toroids in $Z(\eta(X_\tau))$. Evidently, we can apply the same argument we used in Section 2 to conclude that there is an element z in $Z(\eta(X_\tau)) \cap G$ such that $z\eta(T)z^{-1} = T$. Let z^* denote the inner automorphism of A that is effected by z. Then we have $(z^*\eta)(T) = T$, so that $z^*\eta \in W(T, L)$. Moreover, $(z^*\eta)(X_\tau L) = \eta(X_\tau L)$, so that our two quasi-affine structures of $L\backslash G$ are indeed conjugate by an element of $W(T, L)$. Thus *the conjugacy classes of quasi-affine structures of* $L\backslash G$ *under complex analytic automorphisms are in one to one correspondence, via the map of Proposition* 3.1, *with the orbits of the vector space* $\mathrm{Hom}(\mathrm{Hom}(K(T, L), C^*), Q^L)$ *under the linear action of* $W(T, L)$. Moreover, it is clear from the above proof that *the same result holds if the group* W *of all complex analytic automorphisms of* G *is replaced with the group of the inner automorphisms.*

In certain cases, one can show that the orbits for the action of $W(T, L)$ on $\mathrm{Hom}(\mathrm{Hom}(K(T, L), C^*), Q^L)$ are discrete, or even finite. The discussion of this requires some preparation.

Let T^0 be a maximal complex pro-toroid in A that contains T, and let

Z denote the centralizer of T in A. Then we have $\eta(Z)=Z$ for every η in $W(T,L)$, so that both T^0 and $\eta(T^0)$ are maximal complex pro-toroids in Z. Hence there is an element z in $Z \cap G$ such that $z\eta(T^0)z^{-1}=T^0$. Thus, if z^* denotes the inner automorphism effected by z, we have $z^*\eta \in W(T^0)$, where $W(T^0)$ denotes the stabilizer of T^0 in W. Hence $W(T,L) \subset (Z \cap G)^* W(T^0)$. Since the elements of Q and T are left fixed under the action of $(Z \cap G)^*$, this shows that each orbit for the action of $W(T,L)$ on Q^L is contained in an orbit for the action of $W(T^0)$ on Q, and each orbit for the action of $W(T,L)$ on $K(T,L)$ is contained in the canonical image in $(T^0 \cap G)/(T \cap L)$ of an orbit for the action of $W(T^0)$ on $T^0 \cap G$.

The situation with regard to the inner automorphisms is quite simple. The result is the following.

THEOREM 4.1. *The group of automorphisms of*

$$\mathrm{Hom}(\mathrm{Hom}(K(T,L),C^*),Q^L)$$

that corresponds to the group $G^* \cap W(T,L)$ *of the inner automorphisms stabilizing* T *and* L *is finite.*

Proof. Since the elements of Q are fixed under G^*, it suffices to show that $G^* \cap W(T,L)$ induces a finite group of automorphisms on $K(T,L)$. A fortiori, it suffices to show that this group induces a finite group of automorphisms on $T \cap G$. From the above, we have

$$G^* \cap W(T,L) \subset (Z \cap G)^*(G^* \cap W(T^0)).$$

Since $(Z \cap G)^*$ acts trivially on $T \cap G$, it suffices therefore to show that $G^* \cap W(T^0)$ induces a finite group of automorphisms on the larger group $T^0 \cap G$. Thus it suffices to show that $N_G(T^0 \cap G)$ induces a finite group of automorphisms on $T^0 \cap G$, via conjugation.

We have a semidirect product decomposition $G=H \cdot K$, where K is a nucleus of G and H is a reductive subgroup containing $T^0 \cap G$. Now let h be an element of H, let k be an element of K, and suppose that hk normalizes $T^0 \cap G$. Then, for every element x of $T^0 \cap G$, we have $(hk)x(hk)^{-1} \in T^0 \cap G$, whence $kxk^{-1} \in H$, so that $kxk^{-1}x^{-1} \in H \cap K=(1)$. Thus k centralizes $T^0 \cap G$, so that the automorphism effected by hk on $T^0 \cap G$ coincides with the automorphism effected by h. Hence it suffices to show that $N_H(T^0 \cap G)$ induces a finite group of automorphisms on $T^0 \cap G$. We have $H=H'Z(H)$, where $Z(H)$ is the connected component of 1 in the center of H. Since the maximal complex toroid $T^0 \cap G$ in H contains $Z(H)$, by Lemma 4.3 of [3], we have therefore $T^0 \cap G=(T^0 \cap H')Z(H)$. Moreover, $T^0 \cap H'$ is a maximal complex toroid in H', so that its Lie algebra is a Cartan subalgebra of the semi-simple Lie algebra of H'. As is well-known, this implies that the group of automorphisms of $T^0 \cap H'$ that is induced by the stabilizer of $T^0 \cap H'$ in the group of all complex analytic automorphisms of H' is finite. In particular, this proves Theorem 4.1.

For later use, we put some of our information on $T^0 \cap G$ into an invariant form. Note that G' is the semidirect product $H' \cdot N$, where N is the radical of G', which is contained in K. We see easily from this that, actually, $T^0 \cap H' = T^0 \cap G'$. On the other hand, the radical R of G is $Z(H)K$. Hence it is clear that $Z(H)$ is the connected component of 1 in $T^0 \cap R$. Let us write D for $T^0 \cap G$, and, generally, let $P^{\cdot}$ stand for the Lie algebra of the analytic group P. Then it is clear from the above that $D^{\cdot}$ is the direct sum of $D^{\cdot} \cap [G^{\cdot}, G^{\cdot}]$ and $D^{\cdot} \cap R^{\cdot}$. Evidently, each of these two summands is stable under the canonical action of $W(T^0)$. Moreover, $D^{\cdot} \cap [G^{\cdot}, G^{\cdot}]$ is canonically isomorphic with the Cartan subalgebra $(D^{\cdot} + R^{\cdot})/R^{\cdot}$ of the semisimple Lie algebra $G^{\cdot}/R^{\cdot}$, whence we conclude that the group of automorphisms of $D^{\cdot} \cap [G^{\cdot}, G^{\cdot}]$ that is induced by $W(T^0)$ is finite.

Now we are in a position to prove the following lemma.

LEMMA 4.2. *Let G be a complex linear analytic group, D a maximal complex toroid in G, B the maximum nilpotent normal analytic subgroup of G, N the radical of G'. Let $W(D)$ be the stabilizer of D in the group of all complex analytic automorphisms of G. If B/N has a full compact subgroup then the group of automorphisms of* Hom(G, C) *that is induced by $W(D)$ is finite. If the center of G contains no non-trivial complex toroid then the group of automorphisms of D that is induced by $W(D)$ is finite.*

Proof. Let $Z(D)$ denote the centralizer of D in G. By considering the adjoint action of D on $R^{\cdot}$, we see easily that $R^{\cdot} = (Z(D) \cap R)^{\cdot} + N^{\cdot}$. Let $J^{\cdot}$ be a Cartan subalgebra of $(Z(D) \cap R)^{\cdot}$, and let J be the corresponding analytic subgroup of G. Then, since

$$(Z(D) \cap R)^{\cdot} = J^{\cdot} + [(Z(D) \cap R)^{\cdot}, (Z(D) \cap R)^{\cdot}]$$

and $[R^{\cdot}, R^{\cdot}] \subset N^{\cdot}$, we have $R^{\cdot} = J^{\cdot} + N^{\cdot}$. In particular, we have therefore $G = JG'$.

Now let $W(D, J)$ be the stabilizer of J in $W(D)$. Since any two Cartan subalgebras of $(Z(D) \cap R)^{\cdot}$ are conjugate by an inner automorphism effected by an element of $Z(D) \cap N$, we have $W(D) = W(D, J)\alpha(Z(D) \cap N)$, where α is the canonical homomorphism of G into the group of automorphisms of G. Since the elements of Hom(G, C) and of D are left fixed by $\alpha(Z(D) \cap N)$, the automorphism groups induced by $W(D)$ on Hom(G, C) and on D coincide with those induced by $W(D, J)$.

For every linear map γ of $J^{\cdot}$ into C, let $R^{\cdot}{}_{\gamma}$ denote the space of all elements y in $R^{\cdot}$ such that $[x,y]=\gamma(x)y$ for all elements x of $J^{\cdot}$. Evidently, if $R^{\cdot}{}_{\gamma}\neq(0)$ then γ annihilates $B^{\cdot}\cap J^{\cdot}$. Conversely, suppose that x is an element of $J^{\cdot}$ such that $\gamma(x)=0$ whenever $R^{\cdot}{}_{\gamma}\neq(0)$. Then x annihilates each $R^{\cdot}{}_{\gamma}$, and we see from the decomposition of $R^{\cdot}$ into weight spaces with respect to $J^{\cdot}$ that x is therefore nilpotent on $R^{\cdot}$, so that $x\in B^{\cdot}$. Clearly, $W(D,J)$ permutes the finite set of the $R^{\cdot}{}_{\gamma}$'s. Hence there is a subgroup W^* of finite index in $W(D,J)$ that stabilizes each $R^{\cdot}{}_{\gamma}$. Let η be an element of W^*, let x be an element of $J^{\cdot}$ and let y be an element of $R^{\cdot}{}_{\gamma}$. Then, if η^0 denotes the differential of η, we have

$$\gamma(x)\eta^0(y)=\eta^0(\gamma(x)y)=\eta^0([x,y])=[\eta^0(x),\eta^0(y)]=\gamma(\eta^0(x))\eta^0(y).$$

Hence $\gamma(\eta^0(x)-x)=0$ whenever $R^{\cdot}{}_{\gamma}\neq(0)$, so that $\eta^0(x)-x\in B^{\cdot}$. Hence W^* acts trivially on $J/(J\cap B)$.

Now suppose that B/N has a full compact subgroup. Then $\mathrm{Hom}(G,C)$ is trivial on B. Since $G=JG'$, the restriction map $\mathrm{Hom}(G,C)\to\mathrm{Hom}(J,C)$ is injective. Hence we conclude that W^* acts trivially on $\mathrm{Hom}(G,C)$, so that the group of automorphisms of $\mathrm{Hom}(G,C)$ that is induced by $W(D)$ is finite.

Now note that, since $D^{\cdot}\cap R^{\cdot}$ lies in the center of $(Z(D)\cap R)^{\cdot}$, we have $D^{\cdot}\cap R^{\cdot}\subset J^{\cdot}$. Hence we know from the above that W^* acts trivially on $(D^{\cdot}\cap R^{\cdot})/(D^{\cdot}\cap B^{\cdot})$. On the other hand, if x is an element of $D^{\cdot}\cap B^{\cdot}$ then the adjoint action of x on $G^{\cdot}$ is both semisimple and nilpotent, and therefore trivial. Thus $D^{\cdot}\cap B^{\cdot}$ lies in the center of $G^{\cdot}$. But, by Lemma 5.2 of [3], the connected component of 1 in the intersection of D with the center of G is a complex toroid. Hence if the center of G contains no non-trivial complex toroid we must have $D^{\cdot}\cap B^{\cdot}=(0)$, so that W^* acts trivially on $D^{\cdot}\cap R^{\cdot}$. Hence, in this case, $W(D)$ induces a finite group of automorphisms on $D^{\cdot}\cap R^{\cdot}$. We know already that

$$D^{\cdot}=D^{\cdot}\cap[G^{\cdot},G^{\cdot}]+D^{\cdot}\cap R^{\cdot}$$

and that $W(D)$ induces a finite group of automorphisms on $D^{\cdot}\cap[G^{\cdot},G^{\cdot}]$. Hence we conclude that, in the present case, the group of automorphisms of D that is induced by $W(D)$ is finite, which completes the proof of Lemma 4.2.

Now we can easily derive the following result on the orbits we are interested in.

THEOREM 4.3. *If B/N has a full compact subgroup then the orbits for the action of $W(T,L)$ on* $\mathrm{Hom}(\mathrm{Hom}(K(T,L),C^*),Q^L)$ *are discrete. If $B=N$ then these orbits are finite.*

Proof. Let T^0 be a maximal complex pro-toroid in A that contains T. If B/N has a full compact subgroup it follows immediately from the first part of Lemma 4.2, since we may identity Q with $\mathrm{Hom}(G,C)$, that the orbits for the action of $W(T^0)$ on Q are finite. By what we have seen just before the statement of Theorem 4.1, this implies that the orbits for the action of $W(T,L)$ on Q^L are finite. On the other hand, $\mathrm{Hom}(K(T,L),C^*)$ is simply a finitely generated free abelian group whose automorphisms evidently give rise only to discrete orbits in $\mathrm{Hom}(\mathrm{Hom}(K(T,L),C^*),Q^L)$. Hence it is clear that the orbits for the action of $W(T,L)$ on this last space are discrete.

If $B=N$ then the center of G can evidently not contain any non-trivial complex toroid, so that, by Lemma 4.2, $W(T^0)$ induces finite groups of automorphisms of $T^0\cap G$ and of Q. Now it follows from what just precedes the statement of Theorem 4.1 that the orbits for the action of $W(T,L)$ on $K(T,L)$ and on Q^L are finite, which gives the last part of Theorem 4.3.

5. Examples. We give two examples of the determination of the affine structures for a complex analytic group (taking the subgroup trivial) which differ radically with respect to conjugacy.

I. Let G be the group consisting of all triples (a,b,c) of complex numbers, with $a\neq 0$, and having the multiplication $(a,b,c)(a_1,b_1,c_1)=(aa_1,ba_1+b_1,c+c_1)$. Let α, β, γ be the functions on G sending each triple (a,b,c) onto a, b, c, respectively. Put $P=C[\alpha,\alpha',\beta,\gamma]$. It is not difficult to verify that P is a fully stable basic subalgebra of $\boldsymbol{R}(G)$, and thus is not only an affine structure for G but actually the structure of an algebraic linear group. Let X be the fixer of P' in A. The elements of the form $(a,0,0)$ make up a maximal complex toroid C^* of G. Put $T=X\times C^*$. Then T is easily seen to be a maximal complex pro-toroid in A, with which P is associated as $(\boldsymbol{R}(G)^X)'$. The recipe of Theorem 3.2 is easily carried out here, and we find that the affine structures associated with T are the left stable basic subalgebras $C[\alpha q,(\alpha q)',\beta,\gamma]$, where q is an arbitrary element of Q.

We find the conjugacy classes of the affine structures by considering the action of the stabilizer of T in W, according to what we have seen in Section 4. Let η be an element of W that stabilizes T. Then η stabilizes the subgroup C^* of G. It must also stabilize the commutator subgroup of G, which consists of the elements $(1,b,0)$. Finally, η stabilizes also the center of G, which consists of the elements $(1,0,c)$. It follows that $\eta(a,b,c)=(a^e,ub,vc)$, where e is 1 or -1 and u, v are non-zero complex numbers. Applying η to a product, one sees that e must be equal to 1. Hence we have $\alpha\circ\eta=\alpha$, $\alpha'\circ\eta=\alpha'$, $\beta\circ\eta=u\beta$, $\gamma\circ\eta=v\gamma$, so that P is left fixed by every element of the stabilizer of T in W. On the other the elements q of Q

are the functions $\exp(t\gamma)$, where t ranges over C, and $\exp(t\gamma) \circ \eta = \exp(vt\gamma)$. Hence all the structures $C[\alpha q, (\alpha q)', \beta, \gamma]$ in which q is different from 1 make up a single orbit under the action of the stabilizer of T in W. Thus there are exactly two conjugacy classes of affine structures of G. The class of P has P as its only element, and P is distinguished by being the only affine structure of G with which G is an algebraic group.

II. Let G be the group consisting of all 4-tuples (x, a, b, c) of complex numbers, with $x \neq 0$, and having the multiplication

$$(x, a, b, c)(x_1, a_1, b_1, c_1) = (xx_1, a + a_1, b \exp(a_1) + b_1, cx_1 + c_1).$$

Let ρ, α, β, γ be the functions given by $\rho(x, a, b, c) = x$, $\alpha(x, a, b, c) = a$, etc. Put $P = C[\rho, \rho', \alpha, \beta, \gamma]$. This is a left stable basic subalgebra of $\mathbf{R}(G)$, i.e., an affine structure for G. Let X be the fixer of P' in A, write C^* for the subgroup of G consisting of the elements $(x, 0, 0, 0)$, and put $T = X \times C^*$. One sees easily that T is a maximal complex pro-toroid in A, with which P is associated as $(\mathbf{R}(G)^X)'$. The procedure of Theorem 3.2 yields the result that the affine structures associated with T are the left stable basic subalgebras $C[\rho q, (\rho q)', \alpha, \beta, \gamma]$, where q ranges over Q.

Referring to the proof of Lemma 4.2, we see that, in the present case, $W(T)$ acts trivially on $\mathrm{Hom}(\mathrm{Hom}(T \cap G, C^*), Q)$. Indeed, we have $B = N$ and $D^\cdot = D^\cdot \cap R^\cdot$, so that the group W^* acts trivially on D and on $\mathrm{Hom}(G, C)$. The centralizer $Z(D)$ consists of the elements $(x, a, b, 0)$, and J is the subgroup consisting of the elements $(x, a, 0, 0)$. There are only three non-zero $R^\cdot{}_\gamma$'s; namely $J^\cdot$, the Lie algebra of the subgroup $(1, 0, 0, C)$ and the Lie algebra of the subgroup $(1, 0, C, 0)$. One verifies easily that each of these is stable under the action of $W(D, J)$, so that we have $W^* = W(D, J)$. Hence we conclude that $W(T)$ acts trivially on $\mathrm{Hom}(\mathrm{Hom}(T \cap G, C^*), Q)$. Thus the affine structures associated with T are left fixed by the stabilizer of T in W, so that no two of them are in the same conjugacy class. The conjugacy classes of the affine structures of G therefore constitute a 1-dimensional complex vector space, in this case.

REFERENCES.

[1] A. Bialynicki-Birula, G. Hochschild and G. D. Mostow, "Extensions of representations of algebraic linear groups," *American Journal of Mathematics*, vol. 85 (1963), pp. 131-144.

[2] G. Hochschild and G. D. Mostow, "On the algebra of representative functions of an analytic group, II," *American Journal of Mathematics*, vol. 86 (1964), pp. 869-887.

[3] ———, "Affine embeddings of complex analytic homogeneous spaces," *American Journal of Mathematics*, vol. 87 (1965), pp. 807-839.

ISOLATED SINGULARITIES OF SOLUTIONS OF LINEAR ELLIPTIC EQUATIONS.*

By JAMES SERRIN and H. F. WEINBERGER.

---

1. Introduction. It was shown by De Giorgi [1] and Nash [6] that every solution of the self-adjoint uniformly elliptic second order linear equation in divergence form,

$$(a^{ij}u_{x^i})_{x^j}=0, \qquad (a^{ij}=a^{ji}) \tag{1}$$

is Hölder continuous. The Hölder exponent μ depends only on the ellipticity constant λ, where λ is defined by

$$\lambda^{-1}\,|\,\xi\,|^2 \leqq a^{ij}\xi_i\xi_j \leqq \lambda\,|\,\xi\,|^2. \tag{2}$$

More recently Moser [5] showed that positive solutions of (1) obey a Harnack inequality. Using this inequality, Royden [7] and Littman, Stampacchia, and Weinberger [3] independently proved that in $n \geqq 3$ dimensions a solution which approaches plus infinity at the origin is bounded above and below by positive multiples of $|\,x\,|^{2-n}$, and that a fundamental solution actually exhibiting this behavior exists.

Recently one of the present authors [8] has investigated the singularities of solutions of a class of non-linear equations which includes (1) as a special case. It was shown there that every solution u of (1) in the punctured domain $0<|\,x\,|<R$ which is $O(|\,x\,|^{2-n})$ must be of the form $u=ag+w$, where a is a constant, g is any particular solution which behaves like $|\,x\,|^{2-n}$, and w is a regular solution of (1) in $|\,x\,|<R$.

Analogously, it was shown [9] that a bounded solution of (1) in $|\,x\,|>R$ must be of the form $u=u_\infty+O(|\,x\,|^{2-n})$, where u_∞ is a constant. See also [5], p. 589 (footnote). Moreover, the existence of a solution $G(x)$ such that $G(x)\approx|\,x\,|^{2-n}$ at infinity is easily established by the procedures which give the existence of a fundamental singularity at the origin (cf. [8]). Here we use the notation $f(x)\approx|\,x\,|^{2-n}$ to denote the fact that there are positive constants a and b such that $a\,|\,x\,|^{2-n}\leqq f\leqq b\,|\,x\,|^{2-n}$.

Received March 4, 1965.

* This research was supported by the United States Air Force Office of Scientific Research through grant no. AF OSR 883-65 and by the National Science Foundation through grant no. GP3904.

The purpose of the present paper is to refine and generalize these results concerning the behavior of a solution at an isolated singularity. It is shown that there exists a positive constant δ depending only on the ellipticity constant λ, such that if u is a solution for $0 < |x| < R$, and if

$$\limsup |x|^{n-2+\delta} u \leqq 0,$$

$|x| \to 0$, then $u = ag + w$, where $g \approx |x|^{2-n}$ and w is a regular solution in $|x| < R$. In other words, a singular solution other than a fundamental solution must approach both plus and minus infinity at a rate definitely faster than a fundamental solution.

Similarly, there exist a constant κ depending only on λ such that if u is a solution at infinity and $\limsup |x|^{-\kappa} u \leqq 0$, $|x| \to \infty$, then

$$u = u_0 + bG + 0(|x|^{2-n-\nu})$$

where G is a particular solution with $G \approx |x|^{2-n}$, and ν depends only on λ.

Thus any solution which does not approach a limit at infinity must go to both plus and minus infinity at least as fast as $|x|^{\kappa}$ on some sets of points. On the other hand, a solution that goes to a limit more rapidly than $|x|^{2-n}$ must go like $|x|^{2-n-\nu}$. The relation between the constant b and the outflow of u is discussed in Section 7.

In Section 3 we introduce an inversion which is the natural generalization of the Kelvin transformation for the Laplace equation. It takes solutions of equation (1) into solutions of another uniformly elliptic equation of the same form. In particular, solutions in the punctured domain $0 < |x| < R$ go into solutions at infinity, and conversely. This transformation allows us to obtain theorems about behavior at infinity from results concerning behavior at the origin, and vice versa. It is applied to yield our main theorems in Sections 4 and 5.

The above inversion transformation requires the symmetry hypothesis $a^{ij} = a^{ji}$. In Section 6 we present some similar results which can be derived without this assumption. In Section 7 we introduce the concept of outflow and use it to obtain more specific results concerning the behavior of a solution at infinity. The result of Theorem 8, which relates asymptotic behavior, the strong maximum principle at infinity, and outflow should be particularly noted. We conclude the paper with some remarks concerning the solvability of the Dirichlet problem for (1) in exterior domains with prescribed outflow, or with prescribed limiting value at infinity.

For the sake of simplicity we have assumed $n \geqq 3$ in the preceding discussion, and also in the proofs to follow. All our results remain valid when $n=2$, though of course the solutions $g(x)$ and $G(x)$ must in this case behave like $\log|x|$.

By a solution of (1) in a domain Ω we shall always mean a continuous function $u(x)$ having strong derivatives u_x which are locally of class L^2, such that

$$\int a^{ij}u_{x^i}\phi_{x^j}dx=0$$

for any continuously differentiable function ϕ with compact support in Ω. The summation convention will be used throughout, with sums running from 1 to n on each index.

2. An oscillation theorem.

Let $u=u(x)$, $x=(x^1,\cdots,x^n)$, be a solution of (1) in an exterior domain, say $|x|>R$. For any $\sigma>R$ we define $M(\sigma)$ and $m(\sigma)$ to be the maximum and minimum of u on the sphere $|x|=\sigma$. Then the following result holds.

THEOREM 1. *Either u tends to a limit u_∞ as x tends to infinity, or else for all suitably large σ we have the relations*

$$M(\sigma)\geqq A\sigma^\kappa \text{ and } m(\sigma)\leqq -A\sigma^\kappa,$$

where $A>0$ and κ is a positive number depending only on the ellipticity constant λ. If $n\geqq 3$ then the limiting value u_∞ is finite.

Remark. In [5] it is shown that if $M(\sigma)$ is increasing and $m(\sigma)$ is decreasing, then $M(\sigma)-m(\sigma)$ grows at least like a power of σ. The present result provides a strengthening of this conclusion.

Proof of Theorem 1. According to [9], Lemma 11, if u is bounded on one side then it tends to a limit u_∞ as $|x|$ tends to infinity; when $n\geqq 3$ the limit u_∞ is finite, while for $n=2$ it may be infinite. We may thus assume in what follows that u becomes both positively and negatively unbounded as x tends to infinity. By the maximum principle, it is clear that $M(\sigma)$ can have no more than one relative minimum for $\sigma>R$, while similarly $m(\sigma)$ can have no more than one relaltive maximum. Then since $M(\sigma)\rightarrow\infty$ and $m(\sigma)\rightarrow-\infty$, there exists some value σ_0 such that M is increasing and m decreasing for $\sigma>\sigma_0$. It may furthermore be assumed that $M>0$ and $m<0$ for $\sigma>\sigma_0$.

We apply the Harnack principle to the solution $u-m(2\sigma)$ in the annulus $\sigma/2<|x|<2\sigma$, where $\sigma>2\sigma_0$. (Clearly $u-m(2\sigma)>0$ in this region). There results easily

$$M(\sigma)-m(2\sigma)\leqq C(m(\sigma)-m(2\sigma))$$

where $C > 1$ depends only on λ. The same argument applied to $M(2\sigma) - u$ leads to

$$M(2\sigma) - m(\sigma) \leqq C(M(2\sigma) - M(\sigma)).$$

Adding the above two inequalities and transposing, we obtain

$$M(2\sigma) - m(2\sigma) \geqq [M(\sigma) - m(\sigma)](C+1)/(C-1).$$

By iteration, we get

$$M(2^\nu\sigma) - m(2^\nu\sigma) \geqq [M(\sigma) - m(\sigma)][(C+1)/(C-1)]^\nu, \quad \nu = 1, 2, \cdots.$$

We conclude (cf. [5], p. 580) that

$$M(\sigma) - m(\sigma) \geqq K\sigma^\kappa, \qquad \sigma > 2\sigma_0,$$

where

$$\kappa = \log_2[(C+1)/(C-1)].$$

By the first two of the above inequalities,

$$(C-1)m(2\sigma) \leqq -[M(\sigma) - Cm(\sigma)] \leqq -[M(\sigma) - m(\sigma)],$$
$$(C-1)M(2\sigma) \geqq CM(\sigma) - m(\sigma) \geqq M(\sigma) - m(\sigma).$$

Thus $m(\sigma) \leqq -A\sigma^\kappa$ and $M(\sigma) \geqq A\sigma^\kappa$, where $A = K/(C-1)$, and $\sigma > 4\sigma_0$.

The following corollary of Theorem 1 restates the result in a form more useful for our purposes.

COROLLARY. *There exists a constant* $\kappa > 0$, *depending only on the ellipticity constant* λ, *such that if* u *is a solution of* $(a^{ij}u_{x^i})_{x^j} = 0$, *and either*

$$\liminf_{|x|\to\infty} |x|^{-\kappa}u \leqq 0 \quad \textit{or} \quad \limsup_{|x|\to\infty} |x|^{-\kappa}u \geqq 0,$$

then u *approaches a limit at infinity.*

Theorem 1 states that if a solution $u(x)$ of (1) in $n \geqq 3$ dimensions is to go to infinity, it must go at least as fast as $|x|^\kappa$ for some definite κ. In the case of Laplace's equation $\kappa = 1$. The following example shows that, in general, κ depends on λ. The equation

$$\{(1+\alpha^2)^{-\frac{1}{2}}(\delta_{ij} + \alpha^2 x^i x^j/|x|^2)u_{x^i}\}_{x^j} = 0 \tag{3}$$

is uniformly elliptic with the ellipticity constant $\lambda = (1+\alpha^2)^{\frac{1}{2}}$. It has the solution

$$u = x^1|x|^{-q}, \qquad q = \tfrac{1}{2}\{n - [(n-2)^2 + 4(n-1)(1+\alpha^2)^{-1}]^{\frac{1}{2}}\}$$

which goes to infinity at the rate $|x|^{1-q}$. Hence

$$\kappa(\lambda) \leqq \tfrac{1}{2}\{[(n-2)^2 + 4\lambda^{-2}(n-1)]^{\frac{1}{2}} - (n-2)\}.$$

For large λ this number is arbitrarily near zero. Hence we cannot choose κ independent of λ.

If the coefficients a^{ij} tend to a limit at infinity, a simple change of variables reduces the problem to the case where λ is arbitrarily near one. In this case Finn and Gilbarg showed (for $n \leqq 6$) that one can take $\kappa = [n-1]^{\frac{1}{2}} - \tfrac{1}{2}(n-2)$.

3. A Kelvin transformation. In this section we establish a generalization of the Kelvin inversion transformation for Laplace's equation. This transformation allows us to take a uniformly elliptic equation of the form (1) defined in a neighborhood of infinity into another equation of the same form defined in a punctured neighborhood of the origin, and vice versa. We first prove the following lemma:

LEMMA. *Let u and w be two solutions of the elliptic equation*

$$(a^{ij}u_{x^i})_{x^j} = 0$$

in a domain Ω, with $w > 0$. Let $y^k = y^k(x^1, \cdots, x^n)$, $k = 1, \cdots, n$, be a continuously differentiable one-to-one coordinate transformation with non-vanishing Jacobian $J = \det(\partial y^k/\partial x^i)$ and inverse $x^i = x^i(y^1, \cdots, y^n)$. Then the function

$$v(y) = u(x(y))/w(x(y))$$

is a solution of the elliptic equation

$$(\bar{a}^{kl}v_{y^k})_{y^l} = 0 \tag{4}$$

in the image domain Ω', where

$$\bar{a}^{kl} = \frac{w^2}{J} a^{ij} \frac{\partial y^k}{\partial x^i} \frac{\partial y^l}{\partial x^j}. \tag{5}$$

Proof. Let $\phi(y)$ be a continuously differentiable function having compact support in Ω'. Then

$$\int \bar{a}^{kl}\phi_{y^k}v_{y^l}\, dy = \int (w^2/J)a^{ij}(\partial y^k/\partial x^i)(dy^l/\partial x^j)\phi_{y^k}v_{y^l}J\, dx.$$

Since the chain rule and the product rule hold for strong derivatives, we have

$$\begin{aligned}\int \bar{a}^{kl}\phi_{y^k}v_{y^l}\, dy &= \int w^2 a^{ij}\phi_{x_i}(u/w)_{x^j}\, dx \\ &= \int a^{ij}\phi_{x_i}(wu_{x^j} - uw_{x^j})\, dx \\ &= \int a^{ij}[(\phi w)_{x^i}u_{x^j} - (\phi u)_{x^i}w_{x^j}]\, dx.\end{aligned}$$

(The symmetry of a^{ij} was used in deriving the last expression.) Since ϕw and ϕu are L^2 strongly differentiable functions with compact support in Ω and u and w are solutions of (1), it follows that

$$\int \tilde{a}^{kl}\phi_{y^k}v_{y^l}\,dy = 0.$$

Hence v is a solution of (4).

Suppose now that u is a solution of (1) in the punctured domain $0 < |x| < R$, and that the equation is uniformly elliptic there:

$$\lambda^{-1}|\xi|^2 \leqq a^{ij}\xi_i\xi_j \leqq \lambda|\xi|^2 \tag{6}$$

for all real vectors ξ. Then there is a positive solution $g(x)$ in $0 < |x| < R$ such that

$$|x|^{2-n} \leqq g(x) \leqq c|x|^{2-n} \qquad (n \geqq 3). \tag{7}$$

where c is a positive constant depending only on λ. For example, g can be taken to be a multiple of the Green's function for $|x| < R+1$ with singularity at the origin [3], [8].

We let $w(x) = c^{-\frac{1}{2}}g(x)$, and introduce the inversion $y^k = x^k/|x|^2$. This transformation takes the punctured domain into the exterior domain $|y| > R^{-1}$. We have

$$\partial y^k/\partial x^i = |x|^{-2}(\delta_{ik} - 2x^ix^k/|x|^2), \qquad J = |x|^{-2n},$$

whence

$$\tilde{a}^{kl} = c^{-1}(g|x|^{n-2})^2 a^{ij}(\delta_{ik} - 2x^ix^k/|x|^2)(\delta_{jl} - 2x^jx^l/|x|^2).$$

The matrix $\delta_{ik} - 2x^ix^k/|x|^2$ is unitary. Therefore it follows from (6) that

$$(c\lambda)^{-1}|\eta|^2 \leqq \tilde{a}^{kl}\eta_k\eta_l \leqq c\lambda|\eta|^2.$$

Thus the equation (4) satisfied by

$$v(y) = u(y/|y|^2)/g(y/|y|^2)$$

is again uniformly elliptic, and its ellipticity constant $\lambda' = c\lambda$ depends only on the ellipticity constant λ of (1). We have proved

THEOREM 2. *If u is a solution of the uniformly elliptic equation* (1) *in* $0 < |x| < R$, *then*

$$v(y) = u(y/|y|^2)/g(y/|y|^2)$$

is a solution of the uniformly elliptic equation (4) *in* $|y| > R^{-1}$, *and the* *ellipticity constant λ' of* (4) *depends only on the ellipticity constant λ of* (1).

Similarly, on using a solution $G(x)$ of (1) which satisfies

$$|x|^{2-n} \leqq G(x) \leqq C|x|^{2-n} \qquad (n \geqq 3) \tag{8}$$

with C depending only on λ, we find:

THEOREM 3. *If u is a solution of the uniformly elliptic equation* (1) *in $|x| > R$, then*

$$v(y) = u(y/|y|^2)/G(y/|y|^2)$$

is a solution of the uniformly elliptic equation (4) *in $0 < |y| < R^{-1}$, and the ellipticity constant λ' of* (4) *depends only on the ellipticity constant λ of* (1).

Remark. By the same proofs with $w \equiv 1$ we can show that Theorems 2 and 3 still hold in two dimensions. In fact, $\lambda = \lambda'$ in this case.

4. Behavior at an isolated singularity. A solution $g(x)$ of (1) in a punctured neighborhood of the origin will be called a *fundamental solution* if $g \approx |x|^{2-n}$, or $g \approx -\log|x|$ if $n = 2$. In the same way, a solution G at infinity with $G \approx |x|^{2-n}$ (or $G \approx \log|x|$) will also be called a fundamental solution.

THEOREM 4. *There exists a positive number δ depending only on the elliptic constant λ, with the following property: If u is a solution of the self-adjoint uniformly elliptic equation* (1) *in the domain $0 < |x| < R$ and if* $\limsup |x|^{n-2+\delta}u \leqq 0$ (*or* $\liminf |x|^{n-2+\delta}u \geqq 0$), $x \to 0$, *then*

$$u = ag + w \tag{9}$$

where a is a constant, g is any given fundamental solution, and w is a solution of (1) *in the whole disk $|x| < R$.*

Proof. We first prove the theorem for a fixed fundamental solution g satisfying (7). By Theorem 2 the function

$$v(y) = u(y/|y|^2)/g(y/|y|^2)$$

satisfies equation (4), which is uniformly elliptic with ellipticity constant $\lambda' = c\lambda$. By (7) the limit superior of $|y|^{-\delta}v$ has the same sign as that of $|x|^{n-2+\delta}u$, for any $\delta > 0$.

We choose δ to be the constant κ of Theorem 1 corresponding to the ellipticity constant $\lambda' = c\lambda$. Then, by Theorem 1, v must be bounded in a neighborhood of infinity. It follows that $u(x)/g(x)$ is bounded near the origin. Hence by Theorems 14 and 15 of [8] u is of the form (9). Since c depends only on λ and δ depends only on $\lambda' = c\lambda$, δ depends only on λ. Thus the theorem is proved for our particular function g.

Now let $\bar{g}$ be any fundamental solution. By what has already been shown, $\bar{g} = \tilde{a}g + \tilde{w}$ where $\tilde{a} > 0$. Then $g = (\bar{g} - \tilde{w})/\tilde{a}$, and hence for any u satisfying the hypotheses of the theorem we have $u = ag + w = \bar{a}\bar{g} + \bar{w}$. Since we can obviously omit the bars on the right side, this completes the proof.

Remark. For $n = 2$ an application of Theorem 1 to $u(y/|y|^2)$ shows that $u(x)$ is bounded either above or below near the origin. It follows from [8, Theorem 14] that u has the form (9).

Theorem 4 states that a function with an unremovable isolated singularity different from the fundamental solution g must go to both plus and minus infinity at least like g times $|x|^{-\delta}$. In the case of Laplace's equation, $\delta = 1$.

We again examine the equation

$$\{(1 + \alpha^2)^{-\frac{1}{2}}(\delta_{ij} + \alpha^2 x^i x^j / |x|^2) u_{x^i}\}_{x^j} = 0,$$

where α is a constant. As we have seen, $\lambda = (1 + \alpha^2)^{\frac{1}{2}}$. The function

(10) $\quad u = x^1 |x|^{-q}$, where $q = \frac{1}{2}\{n + [(n-2)^2 + 4(n-1)(1+\alpha^2)^{-1}]^{\frac{1}{2}}\}$,

is a solution. Hence

$$\delta(\lambda) \leqq \tfrac{1}{2}\{[(n-2)^2 + 4\lambda^{-2}(n-1)]^{\frac{1}{2}} - (n-2)\}.$$

This number can be made arbitrarily close to zero by choosing λ large enough. Hence it is not possible to choose δ independent of the ellipticity constant λ.

5. Behavior at infinity. We shall now extend the result of Theorem 1 to obtain a sharper asymptotic estimate.

THEOREM 5. *Let $\kappa = \kappa(\lambda)$ be the constant of Theorem 1. Let u be a solution of the self-adjoint uniformly elliptic equation (1) for $|x| > R$. There exists an exponent ν depending only on λ, such that if* $\limsup |x|^{-\kappa} u \leqq 0$ (*or* $\liminf |x|^{-\kappa} u \geqq 0$), $x \to \infty$, *then*

(11) $$u(x) = u_\infty + bG(x) + 0(|x|^{2-n-\nu}),$$

where u_∞ and b are constants, and G is any given fundamental solution at infinity. Moreover the last term satisfies a strong maximum principle at infinity.

(We say that a function u satisfies a strong maximum principle at infinity if u tends to a limit u_∞ as x tends to infinity and $u - u_\infty$ either changes sign or is identically zero in every neighborhood of infinity.)

Proof. We first consider a particular G satisfying (8). By Theorems 1 and 3 the function

$$v(y) = u(y/|y|^2)/G(y/|y|^2)$$

satisfies the hypotheses of Theorem 4. Therefore it is of the form

$$v(y) = ag(y) + w(y) \tag{12}$$

where g is any fundamental solution of the transformed equation (4). By putting $u \equiv 1$ in Theorem 3, we see that $G(y/|y|^2)^{-1}$ is a fundamental solution of (4). Hence

$$v(y) = u_\infty G(y/|y|^2)^{-1} + w(y)$$

where u_∞ is a constant, and $w(y)$ is a solution of (4) for $|y| < R$. Now w is Hölder continuous with exponent ν depending only on the ellipticity constant $\lambda' = c\lambda$ of (4), which, in turn, depends only on λ. Hence,

$$\begin{aligned} v(y) &= u_\infty G(y/|y|^2)^{-1} + w(0) + \{w(y) - w(0)\} \\ &= u_\infty G(y/|y|^2)^{-1} + w(0) + O(|y|^\nu). \end{aligned}$$

Returning to $u = Gv$ yields (11) for our particular G.

Now let $\bar{G}$ be any fundamental solution. By what has already been shown

$$\bar{G} = bG + O(|x|^{2-n-\nu}),$$

and hence for any u satisfying the hypotheses of the theorem

$$u = u_\infty + \bar{b}\bar{G} + O(|x|^{2-n-\nu}).$$

To complete the proof, it is only necessary to observe that the final term is of the form

$$G(x) \cdot (w'(y) - w'(0))$$

for some regular solution $w'(y)$ of (4); since $w'(y) - w'(0)$ satisfies the maximum principle near $y = 0$, the last statement of the theorem is clearly valid.

Remark. For $n = 2$ we find by Theorem 1, 3 and 4 that $u(y/|y|^2) = ag(y) + w(y)$. By Theorem 3 we can let $g(y) = G(y/|y|^2)$. Inverting again yields (11). Of course, u_∞ is no longer the limit of u when $b \neq 0$.

In the special case when $u - u_\infty = o(|x|^{2-n})$, Theorem 5 states that $u - u_\infty = O(|x|^{2-n-\nu})$ and that u satisfies a strong maximum principle at infinity. The latter part of this result is due to Finn and Gilbarg [2, Section 12], and holds even when (2) is replaced by the weaker condition

$$0 < a^{ij}\xi_i\xi_j \leqq \lambda\,|\,\xi\,|^2, \qquad \xi \neq 0.$$

For Laplace's equation, clearly $\nu = 1$. The solution (10) of equation (3) shows, however, that

$$\nu(\lambda) \leqq \tfrac{1}{2}\{[(n-2)^2 + 4\lambda^{-2}(n-1)]^{\frac{1}{2}} - (n-2)\},$$

which approaches zero as $\lambda \to \infty$. Thus ν cannot be chosen independent of λ.

6. Non-self-adjoint equations. In this section we shall obtain some related results concerning the behavior of solutions of (1) in the neighborhood of infinity, without making the restrictive assumption $a^{ij} = a^{ji}$.

We first remark that the Harnack inequality has been proved without the symmetry assumption [8]. Therefore, Theorem 1 holds in the non-self-adjoint case. We must, however, add the inequality $|\,a^{ij}\,| \leqq \lambda$ to the definition of the ellipticity constant λ.

We now investigate the behavior of solutions at infinity.

THEOREM 6. *Let u be a solution of* (1) *in the domain $|\,x\,| > R$, and suppose that* $\limsup |\,x\,|^{-\kappa}u \leqq 0$, *where κ is the constant of Theorem* 1. *Then the representation*

$$u = u_\infty + bG + o(|\,x\,|^{2-n}) \tag{13}$$

is valid, where u_∞ and b are constants, and G is any given fundamental solution at infinity.

Remark. The conclusion of Theorem 6 is weaker than that of Theorem 5 in that the last term is now only known to be $o(|\,x\,|^{2-n})$ rather than $O(|\,x\,|^{2-n-\nu})$. We shall show in the next section that it again satisfies a strong maximum principle at infinity.

Proof of Theorem 6. By Theorem 1, u tends to a finite limit u_∞ as x tends to infinity ($n \geqq 3$). We may assume $u_\infty = 0$ without loss of generality in obtaining the representation (13). Since $G > 0$ there exists a constant A such that

$$|\,u\,| < AG \text{ on } |\,x\,| = 2R. \tag{14}$$

Now consider the set of numbers β such that

$$u - \beta G > 0 \tag{15}$$

in some neighborhood of infinity. By (14) and the maximum principle it is evident that this set contains $\beta = -A$, but does not contain $\beta = A$. Let b denote the least upper bound of the values β in this set, and put

$$w = u - bG.$$

We assert that $w = o(G)$ as x tends to infinity.

Indeed, suppose that the assertion is not true. Then there exists an $\epsilon > 0$ and a sequence of points x_m tending to infinity, such that either

$$w(x_m) > \epsilon G(x_m) \tag{16}$$

or

$$w(x_m) < -\epsilon G(x_m). \tag{17}$$

By the definition of b it is clear that

$$u - (b - \epsilon/2)G > 0 \text{ near infinity,}$$

whence the relation (17) cannot be true. It remains to be shown that (16) also cannot hold.

Let θ be a small positive number which will be determined later, and put

$$w' \equiv u - (b - \epsilon\theta)G.$$

By the definition of b, the function w' is positive in some neighborhood of infinity. Moreover, if (16) holds,

$$w'(x_m) = w(x_m) + \theta\epsilon G(x_m) > (1 + \theta)\epsilon G(x_m).$$

By the Harnack inequality applied to the solution w' in the annulus $|x_m|/2 \leqq |x| \leqq 2|x_m|$, one has for all sufficiently large m,

$$w'(x) > [(1+\theta)\epsilon/C]G(x_m), \qquad |x| = |x_m|, \tag{18}$$

where C is the Harnack constant already used in the proof of Theorem 1. A second application of the Harnack inequality similarly establishes that

$$G(x) < CG(x_m), \qquad |x| = |x_m|. \tag{19}$$

Combining (18) and (19) we have for all sufficiently large m

$$w'(x) > [(1+\theta)\epsilon/C^2]G(x) \text{ on } |x| = |x_m|.$$

Hence by the maximum principle

$$w' > [(1+\theta)\epsilon/C^2]G(x)$$

near infinity. This may be written in the form

$$u - \{b + [(1+\theta)/C^2 - \theta]\epsilon\}G > 0. \tag{20}$$

Now choose $\theta < 1/(C^2-1)$. Then $(1+\theta)C^{-2}-\theta > 0$, and inequality (20) violates the definition of b as the least upper bound of the numbers β such that (15) holds in some neighborhood of infinity. This contradiction establishes the assertion $w = o(G)$.

Since $G \approx |x|^{2-n}$, it follows that

$$u = bG + w$$

where $w = o(|x|^{2-n})$. Since we have assumed that $u_\infty = 0$, the theorem is proved.

Remark. When $n = 2$, we observe that the limit value of u may be infinite, but in any case by Theorem 1 and [9, Theorem 2] we have

$$u = O(\log|x|).$$

Hence the consideration beginning with relation (15) can easily be carried out, with the conclusion that $u = bG + w$, where $w = o(\log|x|)$. It then follows (for example, from [9] or from Lemmas 1 and 2 of Section 7) that w has a finite limit.

7. Outflow. We shall now relate our results to the outflow of the solution u. If u is continuously differentiable, we define the outflow

$$K[u] = \int_{|x|=R} a^{ij}u_{x^i}x^j/R\,dS$$

where $|x| = R$ is any sphere such that u is a solution in $|x| \geqq R$. In general, we define the outflow by the formula

$$K[u] = \int a^{ij}u_{x^i}\psi_{x^j}\,dx$$

where ψ is any strongly differentiable continuous function whose support lies in the set $|x| > R$, and which is identically one in some neighborhood of infinity.

Since the difference of two such functions ψ has compact support, and since u is a solution of (1), it follows that $K[u]$ is independent of the particular ψ which is chosen. Since the support of ψ_x is finite, $K[u]$ is finite. The two definitions for $K[u]$ given above are easily seen to be in agreement when u is smooth.

The following two lemmas relate the outflow to the behavior at infinity. The first lemma states that a solution at infinity which does not satisfy the strong maximum principle there (in particular, a fundamental solution G) has non-zero outflow. The second lemma says that any solution which is $o(G)$ has zero outflow.

LEMMA 1. *Let v be a solution of* (1) *in a domain which contains a neighborhood* $|x| \geqq R$ *of infinity. Let v have a limit v_∞ (possibly $-\infty$) as $|x| \to \infty$, and let $v \geqq v_\infty$. Then either $v \equiv v_\infty$ or $K[v] < 0$.*

Proof. Suppose v is not constant. By the Harnack principle $v > v_\infty$, and hence there are constants A and B such that $v > A > B > v_\infty$ for $|x| = R$. Let

$$\psi = \begin{cases} 0 & \text{when } v > A \\ (A - v)/(A - B) & \text{when } B \leqq v \leqq A \\ 1 & \text{when } v < B. \end{cases}$$

Then by definition

$$K[v] = -(A - B)^{-1} \int_{B \leqq v \leqq A} a^{ij} v_{x^i} v_{x^j} \, dx < 0$$

since v_x cannot vanish almost everywhere on the set $B \leqq v \leqq A$.

LEMMA 2. *Let w be a solution at infinity of the uniformly elliptic equation* (1). *If $w = o(|x|^{2-n})$ (or $o(\log |x|)$ if $n = 2$), then $K[w] = 0$.*

Proof. Let G be a fundamental solution satisfying (8). We observe that $K[u]$ is a linear functional, so that for any constant α

$$K[w + \alpha G] = K[w] + \alpha K[G]. \tag{21}$$

When $\alpha > 0$, $w + \alpha G$ is again a fundamental solution at infinity since $w = o(|x|^{2-n})$. Hence by Lemma 1, $K[w + \alpha G] < 0$ when $\alpha > 0$. Similarly, $K[w + \alpha G] > 0$ when $\alpha < 0$. Therefore we see from (21) that $K[w] = 0$, which is the required conclusion.

When $n = 2$, Lemma 1 shows that $K[G] > 0$ and we must only reverse the inequalities to obtain the same result.

By combining Lemmas 1 and 2 with Theorems 6 and 5, we obtain the following two theorems:

THEOREM 7. *Let u be a solution at infinity of* (1) *with* $\limsup |x|^{-\kappa} u \leqq 0$. *Then there exists a constant u_∞ such that if G is any fundamental solution, then*

$$u = (K[u]/K[G])\, G + u_\infty + o(|x|^{2-n}).$$

If $a^{ij} = a^{ji}$, the last term becomes $O(|x|^{2-n-\nu})$.

THEOREM 8. *The following three statements about a bounded solution u at infinity are equivalent:*

(a) $u - u_\infty = o(|x|^{2-n})$,

(b) *u satisfies a strong maximum principle at infinity,*

(c) $K[u] = 0$.

As we have remarked earlier, the conclusion (a) $\Rightarrow$ (b) is due to Finn and Gilbarg [2].

8. The exterior Dirichlet problem. We conclude the paper with some remarks concerning the solvability of the Dirichlet problem for (1) in exterior domains.

Let D be a given open region containing a neighborhood of infinity and having a smooth internal boundary ∂D. A continuous function $\psi(x)$ is prescribed on ∂D. We wish to find a solution u of (1) in D which continuously takes on the values $\psi(x)$ on the boundary ∂D.

Naturally one must specify some further conditions at infinity in order to guarantee uniqueness of the solution. If $n = 2$ we require that the solution should be bounded in D (by Theorem 7 this involves the requirement $K[u] = 0$). If $n \geqq 3$ we may specify the solution in either of two ways:

1) The solution should tend to an assigned limit u_∞ as x tends to infinity, or

2) The solution is bounded, and $K[u]$ is assigned.

We shall show in these cases that the exterior problem is well set. Moreover, when $n \geqq 3$ all bounded solutions of (1) taking on the boundary values $\psi(x)$ are realized as u_∞ ranges over the real numbers, or as $K[u]$ ranges over the real numbers.

When $n = 2$ the existence of a bounded solution of the Dirichlet problem is easily established by forming sequences of solutions defined in $D \cap \{|x| < m\}$, as in reference [4]. The uniqueness then follows from the strong maximum principle of Theorem 8.

When $n \geqq 3$ one can find a bounded solution $\bar{u}$ exactly as in the case $n = 2$. In order to find a solution with the appropriate limit at infinity, or the appropriate value of $K[u]$ one may add to $\bar{u}$ another solution which is zero on the internal boundary. In particular, it is clear that there exists a solution $H(x)$ in D such that $H = 0$ on D and $H \rightarrow 1$ as $|x| \rightarrow \infty$. Let

$$u = \bar{u} + \alpha H,$$

where α remains to be chosen. Then

$$u_\infty = \bar{u}_\infty + \alpha,$$
$$K[u] = K[\bar{u}] + \alpha K[H].$$

By Lemma 1, $K[H] \neq 0$, so that we can choose α to give any preassigned value to either u_∞ or $K[u]$.

If u_∞ is assigned, uniqueness follows from the ordinary maximum principle. If $K[u]$ is assigned, uniqueness follows from the strong maximum principle of Theorem 8.

Finally, since every bounded solution has a finite limit at infinity and a finite outflow, it is clear that all bounded solutions of the Dirichlet problem are obtained as the limit value u_∞ in 1) ranges over the real numbers, or else as the outflow $K[u]$ in 2) ranges over the real numbers.

UNIVERSITY OF MINNESOTA.

REFERENCES.

[1] E. De Giorgi, "Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari," *Memorie dell' Accademia delle Scienze de Torino*, Ser. 3a, 3 (1957), pp. 25-43

[2] R. Finn and D. Gilbarg, "Three-dimensional subsonic flows and asymptotic estimates for elliptic partial differential equations," *Acta Mathematica*, vol. 98 (1957), pp. 265-296.

[3] W. Littman, G. Stampacchia and H. Weinberger, "Regular points for elliptic equations with discontinuous coefficients," *Annali della Scuola Normale Superiore di Pisa*, Ser. III, 17 (1963), pp. 45-79.

[4] N. Meyers and J. Serrin, "The exterior Dirichlet problem for second order elliptic partial differential equations," *Journal of Mathematics and Mechanics*, vol. 9 (1960), pp. 513-538.

[5] J. Moser, "On Harnack's theorem for elliptic differential equations," *Communications on Pure and Applied Mathematics*, vol. 14 (1961), pp. 577-591.

[6] J. Nash, "Continuity of solutions of parabolic and elliptic equations," *American Journal of Mathematics*, vol. 80 (1958), pp. 931-953.

[7] H. Royden, "The growth of a fundamental solution of an elliptic divergence structure equation," *Studies in Mathematical Analysis and Related Topics; Essays in honor of G. Polya*, pp. 333-340, Stanford University Press, Stanford, 1962.

[8] J. Serrin, "Local behavior of solutions of quasi-linear equations," *Acta Mathematica*, vol. 111 (1964), pp. 247-302.

[9] ———, "Singularities of solutions of nonlinear equations," *Nonlinear Differential Equations (Proceedings of Symposia on Pure Mathematics*, vol. 17), pp. 68-88, American Mathematical Society, Providence, R. I., 1965.

CALCUTTA UNIVERSITY CENTRAL LIBRARY

ON THE CHOW BUNCHES FOR DIFFERENT PROJECTIVE EMBEDDINGS OF A COMPLETE VARIETY.

By William L. Hoyt.

1. Let $f\colon V \to \boldsymbol{P}$ be an embedding, defined over an algebraically closed field k, of a complete irreducible variety V into a coordinatized projective space $\boldsymbol{P}$. For each $r \geqq 0$ and $s > 0$ we define the *Chow bunch* $\mathfrak{C}(V,f,r,s)$ for f to be the set of Chow forms F^{ξ} for all positive $\boldsymbol{P}$-cycles ξ such that $\operatorname{supp}\xi \subset \mathrm{f}(V)$, $\dim\xi = r$, and $\deg\xi = s$. We recall that the Chow form for an irreducible subvariety $W \subset \boldsymbol{P}$ can be characterized as the irreducible form $F^{W}(X_{qi})$, homogeneous in each of $1+\dim W$ sets of $1+\dim\boldsymbol{P}$ indeterminants $\mathfrak{X}_q = \{X_{qi}\}$, $0 \leqq q \leqq \dim W$, $0 \leqq i \leqq \dim\boldsymbol{P}$, such that $F^{W}(a_{qi}) = 0$ for $1+\dim W$ points $a_q = (a_{qi}) \in \boldsymbol{P}$ if and only if $W \cap (\bigcap_{0\leqq q\leqq\dim W} H_q) = \emptyset$, where each H_q is the hyperplane defined by $\sum_i a_{qi}X_i = 0$; that the Chow form for a positive $\boldsymbol{P}$-cycle $\xi = \sum n_j W_j$ is $F^{\xi} = \prod (F^{W_j})^{n_j}$; that $\deg\xi$ is equal to the degree of F^{ξ} in each of the sets $\mathfrak{X}_q$; and that each $\mathfrak{C}(V,f,r,s)$ is a k-closed subset of the projective space $L(\boldsymbol{P},r,s)$ consisting of all forms over the universal domain which are homogeneous of degree s in each of the sets $\mathfrak{X}_q$ ([1]).

If $f'\colon V \to \boldsymbol{P}'$ is an embedding, defined over k, of V into another coordinatized projective space $\boldsymbol{P}'$, then there are natural homeomorphisms in the Zariski topology between the connected components of the Chow bunches for f and f' which make the Chow form for $\sum n_i f'(W_i)$ correspond to the Chow form for $\sum n_i f'(W_i)$. (Specializations of corresponding forms correspond in the same way since the projections pr_1 and pr_2 of $\boldsymbol{P} \times \boldsymbol{P}'$ induce isomorphisms on $(f \times f')(V)$.) These homeomorphisms cannot in general be extended to morphisms, either of corresponding irreducible components viewed as varieties as in [4], or of corresponding connected components viewed as reduced projective schemes over k as in [2]; for Nagata has given an example of a non-normal point on one $\mathfrak{C}(V,f,0,s)$ which corresponds to a normal point on another $\mathfrak{C}(V,f',0,s)$. ([3], Example 1, p. 173.)

Let $\mathfrak{E}$ denote the set of all projective embeddings of V defined over k (we identify isomorphic projective spaces to avoid logical difficulty here),

---

Received October 30, 1963.

and let $Q = \{Q(f)\}$ denote a set of corresponding irreducible or connected components of Chow bunches for the various $f \in \mathfrak{E}$. Following [2] we call a morphism g of $Q(f)$ *integral* if and only if $Q(f)$ has a cover by affine open sets U such that the affine ring for each U is a finite module over the affine ring for $g(U)$ relative to the module structure determined by g. We define the *Segre product* of a pair $f, f' \in \mathfrak{E}$ to be

$$f * f' = S \circ (f \times f') : V \to \boldsymbol{P}^*$$

where $S: \boldsymbol{P} \times \boldsymbol{P}' \to \boldsymbol{P}^*$ is the Segre embedding of the product $\boldsymbol{P} \times \boldsymbol{P}'$ of the ambient spaces for $f(V)$ and $f'(V)$ into a projective space $\boldsymbol{P}^*$ with dimension $(1 + \dim \boldsymbol{P})(1 + \dim \boldsymbol{P}') - 1$ such that $S((a_i) \times (b_j)) = (c_{ij})$ with $c_{ij} = a_i b_j$ for $0 \leqq i \leqq \dim \boldsymbol{P}$ and $0 \leqq j \leqq \dim \boldsymbol{P}'$, and with the lexicographic ordering of the double indices; and we denote by f^m the m-fold Segre product of f. In this paper we prove the following

THEOREM. i) *For each pair $f, f' \in \mathfrak{E}$ there exists an integral morphism*

$$g(f,f') : Q(f * f') \to Q(f)$$

which extends the natural homeomorphism. ii) *If f, f', and $f'' \in \mathfrak{E}$, then*

$$g(f, f' * f'') = g(f,f') \circ g(f * f', f'').$$

iii) *There exists $f_Q \in \mathfrak{E}$ such that $g(f_Q, f)$ is an isomorphism for all $f \in \mathfrak{E}$.*

iv) *If V is nonsingular, then for each $f \in \mathfrak{E}$ there exists $m_{f,Q}$ such that $g(f^m, f')$ is an isomorphism for all $f' \in \mathfrak{E}$ and all $m \geqq m_{f,Q}$.*

(*It is easily seen that f^m in* iv) *can be replaced by the embedding determined by the linear system of hypersurface sections of degree m on $f(V)$.*)

2. Fix f, f' and $f * f' = S \circ (f \times f') \in \mathfrak{E}$ as in § 1; let $L(a)$, $M(b)$ and $N(a,b)$ denote the hyperplanes in $\boldsymbol{P}$, $\boldsymbol{P}'$ and $\boldsymbol{P}^*$ defined by $\sum a_i X_i = 0$ $\sum b_j Y_j = 0$ and $\sum a_i b_j Z_{ij} = 0$, resp.; let

$$F^{\xi}(X_{qi}) \in Q(f) \text{ and } F^{\xi*}(Z_{qij}) \in Q(f * f'),$$

$0 \leqq i \leqq \dim P$, $0 \leqq j \leqq \dim P'$, be Chow forms which correspond under the natural homeomorphism, expressed in terms of convenient indeterminants; and let $r = \dim \xi$. Then for each $(r+1)$-tuple $\prod a_q = \prod (a_{qi}) \in \boldsymbol{P}^{r+1}$, the following assertions are equivalent:

i) $F^{\xi*}(a_{qi} b_{qj}) = 0$ for all $\prod b_q = \prod (b_{qj}) \in \boldsymbol{P}'^{r+1}$;

ii) $\bigcap_{0 \leqq q \leqq r} N(a_q, b_q) \cap \operatorname{supp} \xi^* = \emptyset$ for all $\prod b_q \in \boldsymbol{P}'^{r+1}$;

iii) $\bigcap_{0\leqq q\leqq r}(L(a_q)\times \boldsymbol{P}'\cup \boldsymbol{P}\times M(b_q))\cap \operatorname{supp} S^{-1}(\xi)^*=\emptyset$ for all $\prod b_q\in \boldsymbol{P}'^{r+1}$;

iv) $\bigcap_{0\leqq q\leqq r} L(a_q)\cap \operatorname{supp}\xi=\emptyset$;

v) $F^{\xi}(a_{qi})=0$.

Moreover, if W and W^* are corresponding irreducible components of ξ and ξ^*, then for almost all $2r$-tuples $\prod a_q\times\prod b_q\in \boldsymbol{P}^r\times \boldsymbol{P}'^r$ we have

$$F^{W*}(Z_{0ij}, a_{1i}b_{1j},\cdots, a_{ri}b_{rj})=\prod_n\sum_{ij} d_{ni}e_{nj}Z_{0ij}$$

with distinct points $d_n=(d_{ni})\in \boldsymbol{P}$ and $e_n=(e_{ni})\in \boldsymbol{P}'$; and $\deg W$ is equal to the number of points d_n which are contained in $\bigcap_{1\leqq q\leqq r} L(a_q)$. Therefore, substituting indeterminants $X_{qi}Y_{qj}$ for Z_{qij} in $F^{\xi*}$, we obtain

Lemma 1. *$F^{\xi*}(X_{qi}Y_{qj})=F^{\xi}(X_{qi})H(X_{qi}, Y_{qj})$ where H is a non-zero polynomial which has no factors in the X's alone.*

3. Next for $1\leqq i\leqq t$ let $A_i=\sum_i a_{ij}M_{ij}$ be a linear combination of monomials M_{ij} of fixed degree d_i in a set of indeterminants $\mathfrak{X}$; suppose that not all A_i are zero: let $(A_1,\cdots,A_t)$ denote the point with homogeneous coordinates (a_{ij}) in a projective space L; let U be the set of points $(A_1,\cdots,A_t)\in L$ such that $\gcd\{A_1,\cdots,A_t\}$ has fixed degree d; and let L' denote the space of forms of degree d in $\mathfrak{X}$.

Lemma 2. *U is an irreducible locally closed subset of L, and there is a morphism $g\colon U\to L'$ such that $g(A_1,\cdots,A_t)=\gcd\{A_1,\cdots,A_t\}$ for each point $(A_1,\cdots,A_t)\in U$.*

Proof. Let $G=\gcd\{A_1,\cdots,A_t\}$. Then G has degree d if and only if i) there exist polynomials B_i, $1\leqq i\leqq t$, satisfying $B_i=0$ when $A_i=0$, $\deg A_i-\deg B_i=d$ when $A_i\neq 0$, and $A_iB_j=A_jB_i$ for all i, j; but ii) there exist no polynomials C_i, $1\leqq i\leqq t$, satisfying $C_i=0$ when $A_i=0$, $\deg A_i-\deg C_i>d$ when $A_i\neq 0$, and $A_iC_j=A_jC_i$ for all i, j; moreover, when i) and ii) are satisfied, then iii) the B_i are unique up to a common constant multiple and $A_i=aB_iG$ for some $a\neq 0$. Therefore, U is the intersection of the open subset of L defined by ii) and the irreducible closed subset defined by iii). Furthermore, the relations i) determine a morphism $h\colon U\to U\times L''$, with L'' analogous to L, such that

$$h(A_1,\cdots,A_t)=((A_1,\cdots,A_t)\times(B_1,\cdots,B_t)).$$

(The ratios of the coefficients b_{ij} of the B_i can be expressed rationally in terms of the a_{ij} using a minor of maximum rank from i) viewed as linear equations in the b_{ij}; this defines h locally.) Similarly there is a morphism $h'\colon h(U) \to L'$ defined by the relations iii) such that

$$h'((A_1,\cdots,A_t)\times(B_1,\cdots,B_t)) = \gcd\{A_1,\cdots,A_t\}$$

for each $(A_1,\cdots,A_t)\times(B_1,\cdots,B_t) \in h(U)$. Then $g = h'\circ h$ is the desired morphism.

4. With notation as in § 2, let

$$F^{\xi *}(X_{qi}Y_{qj}) = \sum A_n(X)M_n(Y)$$

with monomials M_n in Y. Then Lemma 1 implies that

$$F^{\xi} = \gcd\{A_1,\cdots,A_t\}.$$

Therefore, there exist L, U, L' as in § 3, $g\colon U \to L'$ as in Lemma 2, and a morphism $g'\colon U' \to L$ defined on an open subset U' of the ambient space for $Q(f*f')$ such that $U' \supset Q(f*f')$, $g'(F^{\xi *}) = (A_1,\cdots,A_t) \in L$ and $g(A_1,\cdots,A_t) = F^{\xi} \in L'$; and $g\circ g'$ induces a morphism

$$g(f,f')\colon Q(f*f') \to Q(f)$$

which extends the natural homeomorphism. Since $g(f,f')$ is bijective and $Q(f)$ and $Q(f*f')$ are complete, the local ring for each point on $Q(f*f')$ is a finite module over the local ring for the corresponding point on $Q(f)$. Hence $g(f,f')$ is integral. Furthermore, $g(f,f')$ induces a birational map on each irreducible component W of $Q(f*f')$. For if η' and η'' are independent generic specializations over k of a P^*-cycle η, then 2η is a specialization over k of $\eta'+\eta''$; hence if $F^{\xi *}$ is generic on W over k, then every multiple factor of $F^{\xi *}$ is rational over k; and it follows that $F^{\xi *}$ and F^{ξ} have $k(F^{\xi *}) = k(F^{\xi})$ as common minimum field of definition over k. Also if f, f' and $f'' \in \mathfrak{E}$, our definitions imply that $(f*f')*f'' = f*(f'*f'')$; and part ii) of the theorem follows from the fact that both $g(f, f'*f'')$ and $g(f,f')\circ g(f*f',f'')$ are defined by solving for gcd's and can be extended to a common morphism defined on an irreducible locally closed set containing $Q((f*f')*f'')$. Consequently, given $f\in\mathfrak{E}$, we can choose $f'\in\mathfrak{E}$ such that the rank of each local ring for $Q(f*f')$ over the corresponding local ring for $Q(f)$ is maximum; and it follows that $g(f*f',f'')$ is an isomorphism for all $f''\in\mathfrak{E}$.

5. To prove the last part of the theorem we need another lemma. Given $t_{ij}\in k$, $0 \leqq i \leqq \dim \boldsymbol{P}$, $0\leqq j \leqq \dim \boldsymbol{P}'$, let T be the linear morphism such that $T(a_i) = (b_j)$ with $b_j = \sum t_{ij}a_i$ for all $(a_i)\in\boldsymbol{P}$ such that not all $\sum_i t_{ij}a_j = 0$.

Suppose now that $f(V)$ is contained in the domain of T and that $f' = T \circ f$. Let $L(\boldsymbol{P}, r, s)$, $L(\boldsymbol{P}', r, s')$ and $L(\boldsymbol{P}^*, r, s^*)$ be the ambient spaces for $Q(f)$, $Q(f')$ and $Q(f * f')$, resp.; let σ^* be the automorphism of $\boldsymbol{P}^*$ which reverses the lexicographic order associated with the Segre embedding S; and let $h(\sigma^*)$ be the natural automorphism of $L(\boldsymbol{P}^*, r, s^*)$ induced by σ^* (so $h(\sigma^*)(F(Z_{qij})) = F(\sigma^{*-1} Z_{qij})$). Our definitions imply that $h(\sigma^*)(Q(f * f')) = Q(f' * f)$.

LEMMA 3. *There exists a linear morphism $h(T)$ from an open subset of $L(\boldsymbol{P}, r, s)$ into $L(\boldsymbol{P}', r, s')$ such that $Q(f)$ is contained in the domain of $h(T)$ and $h(T) \circ g(f, f') = g(f', f) \circ h(\sigma^*)$.*

Proof. If $\boldsymbol{P} = \boldsymbol{P}'$ and T is an automorphism, we define $h(T)$ to be the natural automorphism of $L(\boldsymbol{P}, r, s)$ induced by T. If $\dim \boldsymbol{P} < \dim \boldsymbol{P}'$ and $T(a_i) = (b_j)$ with $b_j = a_j$ for $j \leqq \dim \boldsymbol{P}$ and $b_j = 0$ for $j > \dim \boldsymbol{P}$, we define $h(T)$ to be the linear embedding which sends $F(X_{qi})$ to $F(X'_{qi})$ where $X'_{qi} = Y_{qi}$ for $i \leqq \dim \boldsymbol{P}$. Clearly the lemma holds in both cases.

If $\dim \boldsymbol{P} > \dim \boldsymbol{P}'$ and $T(a_i) = (b_j)$ with $b_j = a_j$ for $j \leqq \dim \boldsymbol{P}'$, let $h(T)$ be the linear morphism such that $h(T)(F(X_{qi})) = F(X''_{qi})$ where $X''_{qi} = Y_{qi}$ for $i \leqq \dim \boldsymbol{P}'$ and $X_{qi} = 0$ for $i > \dim \boldsymbol{P}'$; let $F^{\xi} \in Q(f)$ and $F^{\xi'} \in Q(f')$ be corresponding points; and let $H \subset \boldsymbol{P}$ and $H' \subset \boldsymbol{P}'$ be hyperplanes defined by $\sum u_j X_j = 0$ and $\sum u_j Y_j = 0$ with $(u_j) \in \boldsymbol{P}'$. Since $f(V)$ is contained in the domain of T and V is complete, a point $(b_j) \in \boldsymbol{P}'$ belongs to $H' \cap \operatorname{supp} \xi'$ if and only if there exists $(a_i) \in H \cap \operatorname{supp} \xi$ such that $T(a_i) = (b_j)$; and it follows from the basic properties of Chow forms that $Q(f)$ is contained in the domain of $h(T)$ and that $h(T)(F^{\xi}) = F^{\xi'}$. Hence if $Q(f)$ is irreducible, the lemma holds in this case also.

If $Q(f)$ is reducible, we need some auxiliary maps. Let $\boldsymbol{P}''$ be the ambient space for $(f' * f')(V)$; let T^* be the natural linear morphism induced by T such that $f' * f' = T^* \circ (f * f')$ (so $T^*(c_{ij}) = (d_{mn})$ with $d_{mn} = c_{mn}$ for $m, n \leqq \dim \boldsymbol{P}'$); let σ'' be the automorphism of $\boldsymbol{P}''$ analogous to σ^*; and let $h(T^*)$ and $h(\sigma'')$ be analogues of $h(T)$ and $h(\sigma^*)$. Then $h(\sigma'')$ induces the identity on the subspace of forms symmetric in the double indices m, n, and this space contains $Q(f' * f')$. Furthermore, $h(T) \circ g(f, f') = g(f', f') \circ h(T^*)$, since both morphisms extend to a morphism defined, on a suitable irreducible locally closed set containing $Q(f * f')$, by equating coefficients of monomials not divisible by X_{qi} with $i > \dim \boldsymbol{P}'$ in relations $A_n B_m = A_m B_n$, $A_n = a B_n G$ as in § 3 and solving for the ratios of the coefficients of the B's. Also $g(f', f') \circ h(\sigma'') \circ h(T^*) = g(f', f) \circ h(\sigma^*)$, since both morphisms extend to a morphism defined, on a suitable irreducible locally closed set containing $Q(f*f')$, by writing $F(X_{qi}Y_{qj}) = \sum A'_n(Y)M'_n(X)$ with monomials M'_n in X and considering relations $A'_nB'_m = A'_mB'_n$, $A'_n = aB'_nG'$ as in § 3 for those indices corresponding to monomials M'_n which are not divisible by X_{qi} with $i > \dim \boldsymbol{P}'$. Therefore $h(T) \circ g(f,f') = g(f',f) \circ h(\sigma^*)$ as desired.

In general T is a composite of morphisms T_i of the types considered above, and we define $h(T)$ to be the corresponding composite of the $h(T_i)$.

6. Finally suppose that V is non-singular. Fix $f \in \mathfrak{C}$, choose $f_Q \in \mathfrak{C}$ as in part iii) of the theorem, and let D and D_Q be divisors on V corresponding to hyperplane sections on $f(V)$ and $f_Q(V)$. Then $mD - D_Q$ is ample for large m (very ample in Grothendieck's terminology), and we can find $f' \in \mathfrak{C}$ and D' corresponding to a hyperplane section of $f'(V)$ such that $mD \sim D_Q + D'$. Also we may assume that the linear system of hyperplane sections on $f^m(V)$ is complete. It follows that there is a linear morphism T from an open subset of the ambient space of $f^m(V)$ to the ambient space of $(f_Q * f')(V)$ such that $f^m(V)$ is contained in the domain of T and $f_Q * f' = T \circ f^m$. (For if (u_i) is a homogeneous generic point over k on $f^m(V)$, there are linear combinations $z_j = \sum_i t_{ij}u_i$ such that (z_j) is a homogeneous generic point on $(f_Q * f')(V)$ over k and $k[u_i]$ is integral over $k[z_j]$. But if there were one k-specialization $(u_i) \to (a_i)$ over $(z) \to 0$ with some $a_i \neq 0$, then there would be infinitely many k-specializations $(u_i) \to (ta_i)$ over $(z_j) \to 0$ which is impossible. Hence $f^m(V)$ is contained in the domain of the linear morphism defined by the t_{ij}.) Since we have $h(T) \circ g(f^m, f_Q * f') = g(f_Q * f', f^m)$ by Lemma 3, it follows that the rank of each local ring for $Q(f^m)$ over the corresponding local ring for $Q(f)$ is as large as possible and consequently that $g(f^m, f'')$ is an isomorphism for all $f'' \in \mathfrak{C}$. This completes the proof of the theorem.

REFERENCES.

[1] W. L. Chow and B. L. van der Waerden, "Zur algebraische Geometrie IX," *Mathematische Annalen*, vol. 113 (1937), pp. 692-704.

[2] A. Grothendieck, *Elements de Geometrie Algebrique*, Ch. I, Institut des Haute Études Scientifiques Publications Mathematiques, No. 4, Paris (1960).

[3] M. Nagata, "On the normality of the Chow variety of positive 0-cycles of degree m in an algebraic variety," *Memoirs of the College of Science*, Kyoto, Series A, vol. XXIX (1955), pp. 165-176.

[4] A. Weil, *Foundations of Algebraic Geometry*, Revised edition, American Mathematical Society Colloquium Publications, vol. XXIX, New York (1962).

TRANSCENDENTAL NUMBERS IN THE P-ADIC DOMAIN.

By WILLIAM W. ADAMS.

1. Introduction. The purpose of this paper is to generalize certain results in the theory of transcendental numbers in the complex field to the P-adic field, and to develop the methods necessary for handling this general problem. The first work done in this direction was by Mahler [12] who proved that α^β was transcendental where $\alpha \neq 0, 1$ and β irrational are algebraic numbers, generalizing the proof of Gelfond. Then Veldkamp [16] proved the same result using the proof of Schneider as a guide. Later Gunther [3] generalized a theorem of Schneider (see LeVeque [9], Vol. II, p. 189) from which the transcendence of e^α (α a non-zero algebraic number) followed in addition to the Mahler theorem. He also generalized some theorems of Gelfond on the approximation of certain transcendental numbers by rational numbers [4]. Içen [5, 6] and Gelfond [2] have also done work in this field.

In this paper a general transcendence theorem in the vein of a theorem of Lang [8] will be proved from which the above results on e^α and α^β follow easily (see Sec. 3). Then we will confine our attention to α^β. The first theorem concerns the algebraic independence of $\alpha^\beta, \alpha^{\beta^2}, \cdots, \alpha^{\beta^{r-1}}$ where $r =$ degree of β over $\boldsymbol{Q}$ and is a generalization of a theorem of Gelfond (see Sec. 6). The second gives a transcendence measure for α^β (see Sec. 7) and the method is a generalization of a method of Feldman as applied by Gelfond. Both of the complex theorems may be found in Gelfond ([2], pp. 132-134).

In order to obtain these results certain facts about P-adic analytic functions must be used. This theory will be developed in an appendix. A very useful concept is the Schnirelman integral, an analogue of the complex line integral. With this notion most of the function theoretic results true for complex analytic functions may be obtained. In our context it is used for estimating functions and allows us to follow more closely the analogous function theoretic arguments of the complex case. As the development of this valuable tool is not readily available in the literature, a detailed account of it will be given in the appendix.

In the P-adic theory of transcendental numbers, besides the generalizations

---

Received March 19, 1964.
Revised Received February 4, 1966.

of the methods used in the complex theory, certain other problems arise. The most important of these is that one must work with functions which are defined only in some neighborhood of some point. Even in the complex theory very little is known about locally defined functions. Because of this the upper bound of the constructed approximation functions having a large number of zeros of high order is not as good. This fact appears in all of the proofs making them tighter in the sense that the degree of freedom available in adjusting the various parameters to make the proof work is narrower. Also in all three of the main theorems the effect is felt in a weakening of the theorem from what it was in the complex case. In Sec. 3 we must assume we have an infinite number of points where the functions take on algebraic values, and impose strong growth conditions. In Sec. 6 the result is proved only when the degree of β is $\geqq 4$ (the case for the degree of $\beta = 3$ may be proved by introducing another variable in the approximating function but this is not necessary in the complex case), and in Sec. 6 the transcendence measure is not quite as good. The transcendence of $\wp(\alpha)$ ($\wp =$ Weierstrass *Pe*-function defined by its laurent series about zero), where $\alpha \neq 0$ is algebraic and in the domain of convergence of $\wp$, does not follow from the theorem of Sec. 3 as it does in the complex case precisely because $\wp$ does not satisfy the growth conditions.

Many of the proofs are, in fact, quite analogous to the corresponding proofs in the complex case. When this is so the proof is not given or else it is briefly outlined and the reader is referred to Gelfond [2] or Lang [8] or to my thesis [19].

All of the algebraic number theory used in this paper may be found in Lang [7].

In conclusion I would like to express my indebtedness to Serge Lang for suggesting the topic and for invaluable advice and encouragement.

2. Notation and background.

First we will set up the notation that will be used throughout this paper.

Let P be a positive prime rational integer. Denote by $\boldsymbol{Z}$ the rational integers and by $\boldsymbol{Q}$ the rational numbers. Let $\boldsymbol{Q}_P$ be the completion of $\boldsymbol{Q}$ with respect to the P-adic valuation $|\ |_P$ on $\boldsymbol{Q}$. Let $\boldsymbol{T}_P$ be the completion of the algebraic closure of $\boldsymbol{Q}_P$, and denote its valuation by $|\ |_P$ also. We know that $\boldsymbol{T}_P$ is algebraically closed.

Now on $\boldsymbol{Q}$ (or on the complex numbers) we have the usual archimedian absolute value $|\ |$. If K is a number field, we set $S_\infty =$ the set of all archimedian absolute values on K and set for all $a \in K$, $\| a \| = \max | a |_v$ $(v \in S_\infty)$.

Also if $f(X) = a_0 + a_1X + \cdots + a_nX^n$ is a polynomial with coefficients in K or $\boldsymbol{T}_P$ we denote by $|f|_v = \max|a_i|_v$ $(0 \leqq i \leqq n)$ where v is any absolute value on K or respectively $v = P$. We drop the subscript v if $K = \boldsymbol{Q}$ and v is the usual absolute value. If f has coefficients in K we set $\|f\| = \max\|a_i\|$ $(0 \leqq i \leqq n)$.

By log we will always mean $\log_P$ (log to the base P).

Certain results will now be quoted which will be needed later.

LEMMA 1. *Let* $F_1(X), \cdots, F_n(X)$ *be polynomials with coefficients in* $\boldsymbol{T}_P$ *and set*

$$F(X) = F_1(X) \cdots F_n(X).$$

Then

$$|F|_P = |F_1|_P \cdots |F_n|_P.$$

(This is the Gauss lemma).

LEMMA 2. *Let* $F_1(X), \cdots, F_n(X)$ *be polynomials with complex coefficients and set*

$$F(X) = F_1(X) \cdots F_n(X).$$

Then

$$e^{-\deg F}|F_1| \cdots |F_n| \leqq |F| \leqq e^{\deg F}|F_1| \cdots |F_n|.$$

(See Mahler [13].)

LEMMA 3. *Let* K *be a number field with ring of integers* I_K. *Consider the system of linear equations*

$$\begin{matrix} a_{11}X_1 + \cdots + a_{1n}X_n = 0 \\ \vdots \qquad\qquad \vdots \\ a_{r1}X_1 + \cdots + a_{rn}X_n = 0 \end{matrix}$$

where $n > r$ *and* $a_{ij} \in I_K$ *satisfying* $\|a_{ij}\| \leqq A$ *for* $1 \leqq i \leqq r$ *and* $1 \leqq j \leqq n$. *Then there exists a non-trivial solution* $x_1, \cdots, x_n$ *in* I_K *satisfying*

$$\max_{1 \leqq k \leqq n} \|x_k\| \leqq C_1(nA)^{r/(n-r)}$$

provided $nA > C_2$, *where* C_1 *and* C_2 *are positive constants depending only on* K.

(See Siegel [14].)

LEMMA 4. *Let* $K \subseteq \boldsymbol{T}_P$ *be a number field of degree* N *over the rational numbers. Let* a *be a non-zero algebraic integer in* K. *Then*

$$|a|_P \geqq \|a\|^{-N}.$$

(The proof is a trivial application of the product formula in K.)

Corresponding to the standard inequality for the absolute value of an analytic function with a large number of zeros of a high order we have the following lemma.

LEMMA 5. (Mahler) *Let* $f(z)$ *be a normal function* (*see appendix*) *vanishing at* $x_1, \cdots, x_n$ *in* $\boldsymbol{T}_P$ *of orders* $d_1, \cdots, d_n$ *respectively. Assume* $|x_i|_P \leqq P^{-r}$ *where* r *is a positive integer and let* $d = d_1 + \cdots + d_n$. *Then given any* $x \in \boldsymbol{T}_P$ *such that* $|x|_P \leqq P^{-r}$ *and any* $k = 0, 1, 2, \cdots$ *we have*

$$|f^{(k)}(x)|_P \leqq P^{-r(d-k)}.$$

(This is Theorem 14 of the appendix.)

3. A general transcendence theorem. In [8], Theorem 3, Lang proves a very general transcendence theorem from which one may easily derive most of the standard results on transcendental functions sending algebraic numbers into transcendental numbers. With the analytic tools derived in the appendix, namely the maximum modulus principle, we see that if we define a meromorphic function to be the quotient of two *entire* functions, then Lang's theorem goes through exactly as stated. However, the standard exponential and logarithmic functions are no longer entire but are defined only in some neighborhood of the origin. Also most of the other functions one uses are defined only on some bounded neighborhood, e. g., the Weierstrass $\wp$-function. Thus we see that to get a useful theorem of the type Lang proved, we must deal with *local* functions. The purpose of this section is to prove such a theorem.

THEOREM 1. *Let* $K \subseteq \boldsymbol{T}_P$ *be a number field of degree* N *over* $\boldsymbol{Q}$. *Let* $f_1, \cdots, f_v$ *be P-adic functions meromorphic in some disc* $\mathfrak{D}$. *Let* $c \neq 0$ *be in* $\boldsymbol{T}_P$ *and set* $D = c(d/dz)$. *Consider the ring* $R = K[f_1, \cdots, f_v]$ *and assume* $DR \subseteq R$. *Write* $Df_i = R_i(f_1, \cdots, f_v)$ *where* $R_i \in K[X_1, \cdots, X_v]$. *Let* $\delta = \max \deg R_i$ $(1 \leqq i \leqq v)$ *where* $\deg R_i =$ *maximum of the degrees of the monomials of* R_i. *Assume we have a sequence of distinct points* $z_1, z_2, \cdots$ *in* $\mathfrak{D}$ *isolated from the poles of* $f_1, \cdots, f_v$ *with a limit point in* $\mathfrak{D}$; *such that*

$$f_j(z_i) \in K,\ 1 \leqq j \leqq v \text{ and } i = 1, 2, \cdots.$$

Assume further that we have a positive monotone increasing function σ *on the positive real numbers tending to infinity and*

Case I. $\delta = 1: \sigma(n) = Mn$ *for some* $M > 0$, Case II. $\delta \geqq 2 : \lim_{n \to \infty} \frac{\sigma(n)}{n} = 0$,

and a constant $C > 1$ *depending only on* $R, z_1, \cdots$ *such that*

$$\| f_j(z_n) \| \leqq C^{\sigma(n)} \text{ and } \operatorname{den} f_j(z_n) \leqq C^{\sigma(n)}$$

for $1 \leqq j \leqq v$ *and* $n = 1, 2, \cdots$ (den $=$ *denominator in* K).

Then the transcendence degree of R *over* $K \leqq 1$.

Proof. We may assume the limit point of the sequence $z_1, z_2, \cdots$ is 0. Expand the f_j about 0. Then there are integers a and b such that $P^a f_j(P^b z)$ is a normal function. By passing to a subsequence we may assume that $|z_n|_P \leqq P^{-b-1}$ for all n. Then $P^a f_j(P^b(z_i/P^b)) \in K$ and $|z_i/P^b|_P \leqq P^{-1}$. Thus by considering $P^a f_j(P^b z)$ instead of $f_j(z)$, z_i/P^b instead of z_i, and $P^{\delta(a-1)-b}\gamma D$ instead of D ($\gamma = \operatorname{den} R_i$); we may assume the following:

a) $f_1, \cdots, f_v$ are normal functions (see appendix)

b) $R_1, \cdots, R_v \in I_K[X_1, \cdots, X_v]$ ($I_K =$ integers of K)

c) $|z_i|_P \leqq P^{-1}$, $i = 1, 2, \cdots$.

In order to prove the two cases simultaneously we define $\epsilon = \epsilon(\delta) = 0$ if $\delta = 1$ and $= 1$ if $\delta \geqq 2$.

This proof follows the general outline of Lang's proof in [8].

Let $G(X_1, \cdots, X_v) \in I_K[X_1, \cdots, X_v]$. We will need to calculate an upper bound for $D^k G(f_1, \cdots, f_v)(z_n)$ and $\operatorname{den} D^k G(f_1, \cdots, f_v)(z_n)$. As in [8] set formally $DX_i = R_i(X_1, \cdots, X_v)$ for $1 \leqq i \leqq v$. Also let $B = \max |R_i|$ and $d = \deg G$. Using the notation of [8] we see as there

$$G(X_1, \cdots, X_v) \ll |G| (1 + X_1 + \cdots + X_v)^d$$

for any archimedian absolute value on K. From this we derive inductively

$$\begin{aligned} D^k G(X_1, \cdots, X_v) &\ll \| G \| B^k v^k d^k (k!)^{\epsilon(\delta)} \delta^k (1 + X_1 + \cdots + X_v)^{d+k\delta\epsilon} \\ &\ll \| G \| d^k C_1{}^k k^{\epsilon k} (1 + X_1 + \cdots + X_v)^{d+k\delta\epsilon} \end{aligned}$$

where C_1 depends on the initial data, not on G and k. Thus we obtain

$$\begin{aligned} \| D^k G(f_1, \cdots, f_v)(z_n) \| &\leqq \| G \| d^k C_1{}^k k^{\epsilon k} C^{\sigma(n)(d+k\delta\epsilon)} (v+1)^{d+k\delta\epsilon} \\ &\leqq \| G \| d^k C_1{}^k k^{\epsilon k} C_2{}^{\sigma(n)(d+k\delta\epsilon)}. \end{aligned} \tag{1}$$

Also the denominator is bounded by $C^{\sigma(n)(d+k\delta\epsilon)}$ because the only denominator arises from the $f_i(z_n)$.

Now we will show that any two functions in R are algebraically dependent. It is enough to show this for any $f, g \in \{f_1, \cdots, f_v\}$. Set

$$F(z) = \sum_{i=1}^{s} \sum_{j=1}^{s} a_{ij} f(z)^i g(z)^j$$

where the $a_{ij} \in I_K$ will be chosen not all zero such that $F(z) \equiv 0$. We will let s be arbitrarily large. Set $r_1 = [(\frac{1}{2}) s \log s]$ and $r_2 = [s/\log s]$. We will pick the a_{ij} such that for $0 \leqq k \leqq r_1 - 1$ and $1 \leqq n \leqq r_2$ we have $D^k F(z_n) = 0$. Lemma 3 of Sec. 2 will be applied, as the problem amounts to solving $r_1 r_2 = (\frac{1}{2}) s^2$ equations in s^2 unknowns. So an estimate for the coefficients of the equations and their denominators are needed. The coefficients are of the form $D^k f^i g^j(z_n)$ and by (1) we see

$$\begin{aligned} (2) \qquad \| D^k f^i g^j(z_n) \| &\leqq (i+j)^k k^{\epsilon k} C_1{}^k C_2{}^{\sigma(n)(i+j+k\delta\epsilon)} \\ &\leqq C_3{}^k s^k k^{\epsilon k} C_3{}^{\sigma(n)(s+\epsilon k)}. \end{aligned}$$

Thus with the bound for k and n given above we have

$$\| D^k f^i g^j(z_n) \| \leqq C_4{}^{\sigma(r_2)(s+\epsilon r_1)} C_3{}^{r_1} s^{r_1} r_1{}^{\epsilon r_1}.$$

Similarly

$$\begin{aligned} (3) \qquad \operatorname{den} D^k f^i g^j(z_n) &\leqq C_5{}^{\sigma(n)(s+\epsilon k)} \\ &\leqq C_5{}^{\sigma(r_2)(s+\epsilon r_1)}. \end{aligned}$$

So the a_{ij} exist satisfying

$$(4) \qquad \| a_{ij} \| \leqq C_7{}^{(s+\epsilon r_1)\sigma(r_2)} C_3{}^{r_1} s^{r_1} r_1{}^{\epsilon r_1}.$$

We show that for this choice of a_{ij}, $F(z) \equiv 0$. By Theorem 16 of the appendix it clearly suffices to show that for all positive integers n and k such that $0 \leqq k \leqq r_1 - 1$, $D^k F(z_n) = 0$. If this is not true, choose n least such that $D^k F(z_n) \neq 0$ for some such k. We will estimate $D^k F(z_n)$ from above and below and show these estimates are incompatible.

To calculate a lower bound we calculate an archimedian upper bound and use Lemma 4 of Sec. 2. From (2) and (4) we get

$$\begin{aligned} \| D^k F(z_n) \| &\leqq (\#\,\text{terms}) \max \| a_{ij} \| \, \| D^k f^i g^j(z_n) \|, \qquad 1 \leqq i, j \leqq s \\ &\leqq C_8{}^{\sigma(n)(s+\epsilon r_1)} s^{3r_1} r_1{}^{2\epsilon r_1} \end{aligned}$$

since $n \geqq r_2 + 1$ and $k \leqq r_1 - 1$. Also by (3)

$$\operatorname{den} D^k F(z_n) \leqq C_5{}^{\sigma(n)(s+\epsilon r_1)}.$$

Thus

$$
\begin{aligned}
| D^kF(z_n) |_P &\geqq | (\operatorname{den} D^kF(z_n)) D^kF(z_n) |_P \\
&\geqq C_9{}^{-\sigma(n)(s+\epsilon r_1)} s^{-3Nr_1} r_1{}^{-2N\epsilon r_1}.
\end{aligned} \tag{5}
$$

To calculate an upper bound we apply Lemma 5 of Sec. 2. We have zeros at $z_1, \cdots, z_{n-1}$ of orders at least r_1, so

$$
| D^kF(z_n) |_P = | c |_P{}^k \left| \frac{d^k}{dz^k} F(z_n) \right|_P \leqq C_{10}{}^k P^{-(n-1)r_1+k}. \tag{6}
$$

It is checked easily that (5) and (6) are incompatible for sufficiently large s in the two cases of the theorem. This completes the proof of the theorem.

I note that by essentially the same argument one could show that in Case I it is possible to take any $\sigma(s)$ such that $\lim \sigma(s)/P^{s^{\frac{1}{2}}} = 0$ $(s \to \infty)$ (one must change r_1 and r_2, and modify the last step in the proof). But the solutions to such differential equations are of the form that I can conceive of no case where the stronger result will be necessary.

As a corollary we note (Mahler [12]):

MAHLER'S THEOREM. *Let α and β be algebraic numbers in $\boldsymbol{T}_P$ such that $\alpha \neq 0$, or any root of unity, and β is irrational. Suppose $| \alpha - 1 |_P < 1$ and $| \beta \log \alpha |_P < P^{-1/(P-1)}$. Then α^β is transcendental.*

Proof. One need only set $f_1(z) = e^{z\beta \log \alpha}$, $f_2(z) = e^{z \log \alpha}$, $c = (\log \alpha)^{-1}$, and $z_n = P^r n$ for $n = 1, 2, \cdots$ and r is chosen large (so $| P^r \log \alpha |_P < P^{-1/(P-1)}$).

Similarly if α is algebraic and non-zero and $| \alpha |_P < P^{-1/(P-1)}$ we see that e^α is transcendental. One sets $f_1(z) = z$, $f_2(z) = e^{\alpha z}$, $c = 1$, and $z_n = n$ for $n = 1, 2, \cdots$. The proofs that the above functions f_1, f_2 are algebraically independent proceed exactly as in the complex case. We further note that by defining $\sin z$ and $\cos z$ by their series they also send algebraics into transcendentals. This follows from the same fact about e^z and the formulas still valid in $\boldsymbol{T}_P$:

$$
e^{\sqrt{-1}z} = \cos z + \sqrt{-1} \sin z \text{ and } \cos^2 z + \sin^2 z = 1.
$$

4. Approximating algebraic numbers by integers in a product. The following lemma will be used several times in Secs. 6 and 7. It purpose now is to give a quick proof of a P-adic Liouville theorem.

LEMMA 6. *Let $K \subseteq \boldsymbol{T}_P$ be a number field of degree N over $\boldsymbol{Q}$. Let F and G be polynomials, having no common zeros, with coefficients integers of K. Let $w \in \boldsymbol{T}_P$ and $c = \max\{1, | w |_P\}$. Then*

$$1 \leqq (d_F + d_G)^{N(d_F+d_G)} c^{d_F+d_G} \| F \|^{Nd_G} \| G \|^{Nd_F} \max\{| F(w) |_P, | G(w) |_P\},$$

where $d_F =$ *degree of* F *and* $d_G =$ *degree of* G.

Proof. Let $R = R(F, G)$ denote the resultant of F and G, a non-zero algebraic integer in K by assumption. It is immediate from one of the determinate representations for R that for any archimedean absolute value v

$$| R |_v \leqq (d_F + d_G)^{d_F+d_G} \| F \|^{d_G} \| G \|^{d_F}$$

from which it follows by Lemma 4

$$* \qquad | R |_P \geqq (d_F + d_G)^{-N(d_F+d_G)} \| F \|^{-Nd_G} \| G \|^{-Nd_F}.$$

Now there exist polynomials $A(X)$ and $B(X)$ with rational integer coefficients such that $\deg A < \deg G$ and $\deg B < \deg F$, and $R = AF + BG$ identically in x. Clearly $| A(w) |_P \leqq c^{d_G}$ and $| B(w) |_P \leqq c^{d_F}$. Thus

$$** \qquad | R |_P \leqq c^{d_F+d_G} \max\{| F(w) |_P, | G(w) |_P\}.$$

Combining * and ** we have the desired result.

THEOREM 2. *Let* $\alpha \in \boldsymbol{T}_P$ *be an algebraic number of degree* r *over the rational numbers. Set, for all integers* $s, s_0, \cdots, s_{r-1}, N$, *(where* $s, N > 0$)

$$\beta = 1/s(s_0 + s_1\alpha + \cdots + s_{r-1}\alpha^{r-1}),$$
$$X_N = | \beta |_P | \beta - 1 |_P \cdots | \beta - N |_P.$$

Then if $X_N \neq 0$ *we have*

$$X_N \geqq | N! |_P P^{-c \log^2 M}$$

where c *is a constant depending only on* α *and* $M = \max\{N, s, s_0, \cdots, s_{r-1}\}$.

Proof. The proof will be broken up into two parts, depending on whether β is in $\boldsymbol{Q}_P$ or not.

I). So we assume $\beta \in \boldsymbol{Q}_P$. If $| \beta |_P > 1$, then $X_N > 1$ and we are done. So we assume $| \beta |_P \leqq 1$ and write $\beta = \sum_{k=0}^{v} a_k P^{v_k}$ where v *is* any non-negative integer or ∞, $0 < a_k < P$, and $0 \leqq v_0 < v_1 < v_2 < \cdots$. Then if $A_n = \sum_{k=0}^{n} a_k P^{v_k}$ we see $| \alpha - A_{m-1} |_P = P^{-v_m}$ (set $A_{-1} = 0$, $v_{-1} = 0$). Also if $A_{m-1} < s < A_m$ then $| \alpha - s |_P = | A_{m-1} - s |_P$. Thus choosing n such that $A_{n-1} \leqq N < A_n$ we have

$$X_N = \left(\prod_{i=0}^{n-1} | (A_i - 1 - A_{i-1})! |_P P^{-v_i}\right) | (N - A_{n-1})! |_P^{-v_n} \geqq | N! |_P P^{-\sum_{k=0}^{n} v_k}.$$

Now $v_0 + v_1 + \cdots + v_n \leqq v_n^2$ so we must estimate v_n. To this end let $h(X) \in \boldsymbol{Z}[X]$ be of degree $r-1$ and $f(X) =$ minimal polynomial of α with relatively prime coefficients. Then h and f have no common roots and so we obtain from Lemma 6 that

$$| h(\alpha) |_P \geqq c_1 / | h |^r$$

where c_1 depends only on α. In our case we noted that $| \beta - A_{n-1} |_P = P^{-v_n}$ and so by the above

$$\begin{aligned} P^{-v_n} = | \beta - A_{n-1} |_P \geqq | s\beta - sA_{n-1} |_P &\geqq c_1 (\max\{sN, s_0, \cdots, s_{r-1}\})^{-r} \\ &\geqq c_1 M^{-2r}, \end{aligned}$$

from which it follows immediately $v_n \leqq c_2 \log M$.

II). Let $K = \boldsymbol{Q}(\alpha)$. The valuation $| \ |_P$ on K comes from some prime ideal $\mathfrak{g}$ of K lying over (P). Let $K_{\mathfrak{g}}$ denote the completion of K with respect to $\mathfrak{g}$. There exists a basis of $K_{\mathfrak{g}}$ over $\boldsymbol{Q}_P$, $x_1 = 1, x_2, \cdots, x_n$, such that the x_j are algebraic and $| a_1 x_1 + \cdots + a_n x_n |_P = \max | a_i x_i |_P$ $(1 \leqq i \leqq n)$. Then one may check that we can write $\beta = \beta_1 x_1 + \cdots + \beta_n x_n$ where $\beta_i \in \boldsymbol{Q}_P$ and $\beta_i = 1/g (g_{i0} + \cdots + g_{i,m-1} \gamma^{m-1})$ where γ is algebraic of degree m and depends on α, and g, g_{ij} are integers satisfying $| g |, | g_{ij} | \leqq c_3 M$ for all i, j.

We have now again two cases.

Case a. β_1 is not an integer satisfying $0 \leqq \beta_1 \leqq N$. Then

$$\begin{aligned} X_N = | \beta |_P \cdots | \beta - N |_P &\geqq | \beta_1 |_P | \beta_1 - 1 |_P \cdots | \beta_1 - N |_P \\ &\geqq | N! |_P P^{-c_4 \log^2 M} \end{aligned}$$

by I.

Case b. β_1 is an integer satisfying $0 \leqq \beta_1 \leqq N$. Since $X_N \neq 0$ there is another coordinate of β say $\beta_2 \neq 0$. Then as in part I of the proof

$$| \beta_2 |_P \geqq | g\beta_2 |_P \geqq c_6 (\max_{0 \leqq k \leqq m-1} | g_{2k} |)^{-m} \geqq c_7 / M^m.$$

Thus

$$\begin{aligned} X_N &= | \beta |_P \cdots | \beta - \beta_1 + 1 |_P | \beta - \beta_1 |_P | \beta - \beta_1 - 1 |_P \cdots | \beta - N |_P \\ &\geqq | \beta_1 ! |_P | \beta_2 |_P | (N - \beta_1) ! |_P \geqq | N! |_P c_8 M^{-m} \geqq | N! |_P P^{-c_9 \log M}. \end{aligned}$$

This completes the proof of the theorem.

5. Some lemmas and general comments.

LEMMA 7. *Let F be a polynomial with rational integer coefficients and let $w \in \boldsymbol{T}_P$. Assume $| F | \geqq C$ and $d \geqq C$ (where C is a positive constant depending only on w) and $d \geqq \deg F$. Assume further*

$$|F(w)|_P < H^{-\lambda d}$$

where H, λ are real numbers, $\lambda > 6$, and H satisfies $d < \ln H$ and $|F| \leqq H$. Then there exists a polynomial G in $\mathbf{Z}[X]$ dividing F which is a power of an irreducible polynomial in $\mathbf{Z}[X]$, such that $|G(w)|_P < H^{-(\lambda-6)d}$.

Proof. The proof parallels the proof of the similar result in [2] on page 147. Where Gelfond quotes Lemma V, use Lemma 6 of the preceding section.

LEMMA 8. *Let $\tau_1, \cdots, \tau_n$ be distinct elements of $\boldsymbol{T}_P$ such that $|\tau_i|_P < P^{-1/(P-1)}$ for $1 \leqq i \leqq n$. Let*

$$F(t) = \sum_{i=1}^{n} a_i e^{\tau_i t}$$

where the a_i are also in $\boldsymbol{T}_P$. Set

$$\frac{(z-\tau_1)\cdots\widehat{(z-\tau_k)}\cdots(z-\tau_n)}{(\tau_k-\tau_1)\cdots\widehat{(\tau_k-\tau_k)}\cdots(\tau_k-\tau_n)} = \sum_{j=0}^{n-1} b_{jk} z^j$$

for $1 \leqq k \leqq n$. Then

$$a_k = \sum_{j=0}^{n-1} b_{jk} D^j F(0) \quad \textit{where } D = \frac{d}{dt}.$$

Proof. It is immediate from the above expression defining the b_{jk} that $\sum_{j=0}^{n-1} b_{jk}\tau_i{}^j = \delta_{ki}$ (δ_{ki} = Kronecker delta). Thus

$$\sum_{j=0}^{n-1} b_{jk} D^j F(0) = \sum_{j=0}^{n-1} b_{jk} \sum_{i=1}^{n} a_i \tau_i{}^j = a_k$$

as desired.

The application for this lemma is that if $F \neq 0$ then the order of the largest zero F may have at zero is $n-1$.

LEMMA 9. *Let the notation be as in Lemma 8. Assume further γ, $\kappa \in \boldsymbol{T}_P$ such that $\kappa \neq 0$ and γ is an algebraic irrational of degree r. Let $1 \leqq r_0 \leqq r$. Suppose the τ_i range through all possible $\kappa(i \cdot \gamma)$ where $i \cdot \gamma = i_1 + i_2\gamma + \cdots + i_{r_0}\gamma^{r_0-1}$ and the i's range through $0 \leqq i_j \leqq q_j$. Then*

$$|b_{jk}|_P \leqq P^{cn}$$

where we note that $n = (q_1+1)\cdots(q_{r_0}+1)$ (c does not depend on the q's).

Proof. Compute:

$$\prod_{\substack{j=1\\ j\neq k}}^{n} | \tau_k - \tau_j |_P = | \kappa |_P^{n-1} \prod_{\substack{i_1=0\\ (k_1-i_1)^2+\cdots+(k_{r_0}-i_{r_0})^2\neq 0}}^{q_1} \cdots \prod_{i_{r_0}=0}^{q_{r_0}} | (i_1 - k_1) + \cdots + (i_{r_0} - k_{r_0})\gamma^{r_0-1} |_P$$

$$= | \kappa |_P^{n-1} \Big(\prod_{\substack{i_1=1\\ i_1\neq k_1}}^{q_1} | i_1 - k_1 |_P \Big) \prod_{\substack{i_2=0\\ (k_2-i_2)^2+\cdots+(k_{r_0}-i_{r_0})^2\neq 0}}^{q_2} \cdots \prod_{i_{r_0}=0}^{q_{r_0}} \Big(\prod_{a=-k_1}^{q_1+(-k_1)} | (i_2 - k_2)\gamma + \cdots$$

$$+ (i_{r_0} - k_{r_0})\gamma^{r_0-1} - a |_P \Big)$$

$$\geqq | \kappa |_P^{n-1} | q_1 ! |_P \big(| k_1 ! |_P P^{-c \log^2 q_1} | (q_1 - k_1) ! |_P P^{-c \log^2 q_1} \big)^{(q_2+1)\cdots(q_{r_0}+1)}$$

by Theorem 2 of the preceding section,

$$\geqq | \kappa |_P^{n-1} P^{-q_1/P-1} P^{-n/P-1} P^{-c(q_2+1)\cdots(q_r+1)\log^2 q_1} \geqq P^{-c_2 n}$$

since $| m ! |_P \geqq P^{-m/P-1}$. Moreover the numerator clearly has coefficients $\leqq P^{c_3 n}$ and so the desired result follows.

We now know that α^β is transcendental (notation is as before). As an extension of this result other questions may be asked.

First suppose that the degree of β over $\boldsymbol{Q}$ is $r \geqq 3$. Then what is the transcendence degree of $\alpha^\beta, \alpha^{\beta^2}, \cdots, \alpha^{\beta^{r-1}}$? In Sec. 6 I will prove it is $\geqq 2$ (for $r \geqq 4$; see the comment following the proof). The general conjecture would seem to be that it is $r-1$.

The second question would be in the direction of a transcendence measure for α^β. That is, we know that if $R(X)$ is a non-zero polynomial with integer coefficients, then $R(\alpha^\beta) \neq 0$. Then we would like to know how small it may be as a function of $| R |$ and $\deg R$. In Sec. 7 a transcendence measure is given. It is certainly not the best possible, as one would expect the answer to be something like

$$| R(\alpha^\beta) |_P \geqq | R |^{-(\deg R+1+\epsilon)}$$

for all $\epsilon > 0$. The reason for this is the P-adic version of the Khintchine theorem (Lutz [11] p. 96) which says that all points not satisfying the above have measure zero. To see the most one can hope for I note Theorem 3 below. One sees that the above conjecture in $\boldsymbol{Q}_P$ is precisely what is possible considering this theorem as it guarantees the existence of polynomials R with arbitrarily large absolute value and degree such that

$$| R(\alpha^\beta) |_P < C | R |^{-(\deg R+1)}.$$

THEOREM 3. *Let K be a number field and let $K_{\mathfrak{g}}$ be its completion with respect to some prime $\mathfrak{g}$ lying over P. Let $N_{\mathfrak{g}} = [K_{\mathfrak{g}} : \boldsymbol{Q}_P]$. Let $\alpha_1 \cdots, \alpha_d$ lie in K such that $| \alpha_i |_{\mathfrak{g}} \leqq 1$ for $1 \leqq i \leqq d$. Then for all ν, a* *positive integer, sufficiently large, there exist* $\nu_1, \cdots, \nu_d$, *integers not all zero satisfying* $|\nu_i| \leqq \nu$ $(1 \leqq i \leqq d)$, *such that*

$$|\nu_1\alpha_1 + \cdots + \nu_d\alpha_d|_{\mathfrak{g}} < \frac{C}{\nu^{d/N_{\mathfrak{g}}}},\ C \text{ depends only on } K.$$

Proof. This is an application of the pigeon hole principle. (See Mahler [12] for a similar idea.)

6. On the transcendence degree of $\alpha^{\beta}, \alpha^{\beta^2}, \cdots, \alpha^{\beta^{r-1}}$.

LEMMA 10. *Let* $\sigma(x)$ *be a continuous, strictly monotone, increasing, real-valued function tending to infinity with* x *such that there exists a constant* $a_0 > 1$ *such that* $\sigma(N+1) < a_0\sigma(N)$ *for all* $N > N_0$, N_0 *some positive integer. Let* $w \in \boldsymbol{T}_P$. *Assume that for all integers* $N > N_0$ *we have a polynomial* F_N *with integer coefficients such that*

$$|F_N(w)|_P < P^{-\sigma^2(N)C}$$

where $C = 500a_0^2$, *and*

$$\max\{\deg F_N, \log|F_N|\} \leqq \sigma(N).$$

Then w *is algebraic.*

Proof. This lemma is analogous to Lemma VII, p. 148 in [2] (see also Lang [18], p. 189 for an exposition of the proof). Lemma 6 is used here also.

THEOREM 4. *Let* α *and* β *be algebraic numbers in* $\boldsymbol{T}_P$ *such that* α *is non-zero and is not a root of unity and* β *has degree* $r \geqq 4$. *Suppose* $|\alpha - 1|_P < 1$ *and* $|\beta^k \log \alpha|_P < P^{-1/(P-1)}$, $1 \leqq k \leqq r$. *Then the transcendence degree of* $(\alpha^{\beta}, \cdots, \alpha^{\beta^{r-1}})$ *over* $\boldsymbol{Q}$ *is* $\geqq 2$.

Proof. We follow the proof of Theorem I, p. 155, Gelfond [2].

Let $K = \boldsymbol{Q}(\alpha, \beta, \alpha^{\beta}, \cdots, \alpha^{\beta^{r-1}})$. We assume the theorem is false and hence that K is an algebraic extension of $\boldsymbol{Q}(w)$ where $w = \alpha^{\beta}$. We will show that w is in fact alegbraic giving us the desired contradiction.

By considering some integer multiple of β we may assume β is an algebraic integer. By considering instead of α, α^{P^b} for some large integer b we may assume all functions we deal with are normal; we relabel α^{P^b} as α. Let $D = (\log \alpha)^{-1} \frac{d}{dt}$Denote. $(i \cdot \beta) = i_1 + i_2\beta + \cdots + i_r\beta^{r-1}$ where $(i) = (i_1, \cdots, i_r)$ denotes a vector of non-negative integers.

We set

$$F(t) = \sum_{(i)=(0)}^{(q)} \phi_{(i)}(w)\alpha^{(i\cdot\beta)t}$$

for all sufficiently large integers q where $(q) = \underbrace{(q, \cdots, q)}_{r \text{ times}}$. The $\phi_{(i)}$ are chosen in $\mathbf{Z}[X]$ not all zero such that

$$\deg \phi_{(i)} \leqq q^{r/2}(\log q)^{-2},$$

and

$$\log | \phi_{(i)} | \leqq cq^{r/2}(\log q)^{-2}$$

(c a constant). We want $D^k F(\lambda) = 0$ for all k such that

$$0 \leqq k < c_1 q^{r/2}(\log q)^{-3}$$

where c_1 will be chosen small, and for all $\lambda = P(i \cdot \beta)$ where

$$0 \leqq i_\nu \leqq q^{1/2}(\log q)^{3/r}.$$

We apply Lemma 3 of Sec. 2 and obtain the $\phi_{(i)}$ as desired.

The next step is to let s be the least integer such that for some λ, $D^s F(\lambda) \neq 0$. We will find an upper bound for s. For this we may divide out by the $\phi_{(i)}(w)$ that is largest and then obtain by Lemma 8 (denote by F_1 this new normal function)

$$1 = \sum_{u=0}^{N-1} b_u D^u F_1(0) \text{ where } N = (q+1)^r. \tag{1}$$

Then by Lemma 9

$$| b_u |_P \leqq P^{c_2 N}.$$

Moreover by Lemma 5 we have

$$| D^u F_1(0) |_P \leqq P^{-sm+u}$$

where $m = q^{r/2}(\log q)^3$, the number of λ's. This yields by (1)

$$1 \leqq P^{c_2 N - sm + N}$$

or

$$s \leqq c_3 N/m \leqq c_4 q^{r/2}(\log q)^{-3}.$$

We will now show that the hypothesis of Lemma 10 is satisfied. Set $\sigma(q) = q^{r/2}(\log q)^{-1}$. Let s be the first integer such that $D^s F(\lambda) \neq 0$. Assume now that $K = \boldsymbol{Q}(w, y)$ such that y is algebraic over $\boldsymbol{Q}(w)$. Then $D^s F(\lambda) = \bar{F}(w, y)/E(w)$ where degree $\bar{F}$ in y is bounded independent of q. Let $F_q(w)$ be the resultant of $\bar{F}(w, y)$ and $f(w, y)$, the minimal polynomial of y over $\boldsymbol{Q}(w)$. Then

$$| F_q(w) |_P \leqq | \bar{F}(w, y) |_P \leqq | D^s F(\lambda) |_P \leqq P^{-(s-1)m} \leqq P^{-c_5 q^r} \leqq P^{-\sigma^2(q) \log q}$$

applying Lemma 5 again. Moreover $\max(\deg F_q, \log |F_q|) \leqq \sigma(q)$ may be checked. This shows that w is algebraic, a contradiction. Hence the theorem is proved.

It is possible to prove the theorem in the case that $r = 3$ also. One must then define the approximating function F by

$$F(t) = \sum_{j=0}^{p} \sum_{(i)=(0)}^{(q)} \phi_{(i)}(w) t^j \alpha^{(i \cdot \beta)t}$$

where p is defined appropriately in terms of q. This is considerably more complicated and does not include anything new in the P-adic case, so I do not include it here.

7. A transcendence measure for α^β.

LEMMA 11. *Let $K \subseteq \boldsymbol{T}_P$ be a number field of degree ν over the rational numbers, and let I_K denote its ring of integers. Let $w \in \boldsymbol{T}_P$ and $G(X) \in \mathbf{Z}[X]$, irreducible in $\mathbf{Z}[X]$, such that*

$$|G(w)|_P \leqq H^{-1}$$

where $H \geqq 1$. Then there exists a factor $T(X) \in I_K[X]$ of $G(X)$, irreducible in $I_K[X]$, such that

$$|T(w)|_P < H^{-1/\nu} \text{ and } \| T \| \leqq |G| e^{\deg G}.$$

Proof. This result is an analogue of Lemma VIII, p. 150 of Gelfond [2]. We let $A =$ leading coefficient of G and write $G(X) = AT_1(X) \cdots T_m(X)$ as a product of irreducible monic polynomials in $K[X]$. Then $m \leqq \nu$. Also we may assume that the coefficients of G are relatively prime. Now A is a denominator for each $T_k(X)$: We show $AT_k(X) \in I_K[X]$ by showing that for all primes $\mathfrak{g}$ in I_k, $|AT_k|_\mathfrak{g} \leqq 1$. T_K being monic implies $|T_K|_\mathfrak{g} \geqq 1$ and so by Lemma 1

$$1 = |G|_\mathfrak{g} = |A|_\mathfrak{g} |T_1|_\mathfrak{g} \cdots |T_m|_\mathfrak{g} \geqq |AT_k|_\mathfrak{g}$$

as desired. Now it is easy to check that for some k, $T(X) = AT_k(X)$ satisfies our reguirements (use Lemma 2 to show $\| T \| \leqq e^{\deg G} |G|$).

THEOREM 5. *Let α and β be algebraic numbers in $\boldsymbol{T}_P$ such that α is non-zero and is not a root of unity and β is irrational. Suppose $|\alpha - 1|_P < 1$ and $|\beta \log \alpha|_P < P^{-1/(P-1)}$. Then for all polynomials F with rational integral coefficients such that $|F| > C_0$ (C_0 constant)*

$$|F(\alpha^\beta)|_P > P^{-(\deg F)^3 (\deg F + \log |F|) \log^{2+\epsilon} (\deg F + \log |F|)}$$

for all $\epsilon > 0$.

Proof. This proof is analogous to the proof in Gelfond [2], pp. 162-169. Let $w = \alpha^{\beta}$, so $| w |_P = 1$. The proof proceeds in a number of steps which are outlined below.

1) We construct the approximating function. Take N, q large positive integers such that $N \geqq q \geqq \log^{1+\epsilon} N$ where $\epsilon > 0$ will be chosen small. Set

$$f(z) = \sum_{i=0}^{q} \sum_{j=0}^{N} A_{ij}(w) \alpha^{P^b(i\beta+j)z}$$

where b is a sufficiently large integer and P^b is carried along to guarantee the functions we deal with are normal and so should be ignored. The $A_{ij}(w)$ will be polynomials not all zero with rational integer coefficients and will be chosen by Lemma 3 of Sec. 2 so that we have

$$\deg A_{ij} \leqq p = [q^{\frac{3}{2}} \log^{\frac{1}{2}} N]$$

$$| A_{ij} | \leqq P^{c_1 N q^{\frac{1}{2}} \log^{\frac{1}{2}} N}$$

and

$$D^s f(tP^2) = 0$$

for all integers s, t such that

$$0 \leqq s \leqq r_1 = [N q^{\frac{1}{2}} \log^{-\frac{1}{2}} N]$$

and

$$0 \leqq t \leqq r = [(1/6\nu P^{b+2}) q^{\frac{1}{2}} \log^{\frac{1}{2}} N].$$

We see by Lemma 7 of Sec. 5 that we may assume the polynomials A_{ij} are relatively prime.

2) Now we may estimate f by Lemma 5 of Sec. 2. Then we see that if we let s, t range through the integers

$$0 \leqq s \leqq \bar{r}_1 = [\frac{4}{\epsilon} N q^{\frac{1}{2}} \log^{-\frac{1}{2}} N]$$

$$0 \leqq t \leqq \bar{r} = [q^{\frac{1}{2}} \log^{\frac{1}{2}} N]$$

we have

$$| f^{(s)}(tP^2) |_P \leqq P^{-c_2 N q}, \; c_2 < 1/(6\nu P^{b+2}).$$

3) We note that we may view

$$L_{st}(w) = (\log \alpha)^{-s} d^{s+NtP^{b+2}} f^{(s)}(tP^2)$$

as a polynomial in w of degree $\leqq 2p$ and with coefficients in I_K satisfying

$$\| L_{st} \| \leqq P^{c_3 N q^{\frac{1}{2}} \log^{\frac{1}{2}} N}$$

(Gelfond [2], p. 165) and

$$| L_{st}(w) |_P \leqq P^{-c_4 N q}$$

($d =$ denominator for α and β).

4) We now take a polynomial $G(X) \in \mathbf{Z}[X]$, and for now assume it is irreducible. We will assume it satisfies certain conditions which amount to contradicting the theorem, and deduce a contradiction. The conditions are, with $\delta > 0$ a small real number:

a) $| G(w) |_P < P^{-Nq \log^{3\delta} \log q}$
b) $\deg G < q^{\frac{1}{2}} (\log^{-\frac{1}{2}} N) \log^{-3\delta} \log q$
c) $\log | G | < N q^{-\frac{1}{2}} (\log^{-\frac{1}{2}} N) \log^{-3\delta} \log q.$

5) We obtain immediately from Lemma 11 that there exists a polynomial T with coefficients in I_K which is an irreducible factor of G and satisfies:

a) $| T(w) |_P < P^{-Nq \log^{2\delta} \log q}$
b) $\deg T < q^{\frac{1}{2}} (\log^{-\frac{1}{2}} N) \log^{-2\delta} \log q$
c) $\log \| T \| < N q^{-\frac{1}{2}} (\log^{-\frac{1}{2}} N) \log^{-2\delta} \log q.$

6) It is now an easy application of Lemma 6 of Sec. 4 to show that $T(X)$ divides $L_{st}(X)$ for all s, t in our range (p. 166, Gelfond [2]).

7) The result in 6 can be used to give a better bound for $| f^{(s)}(tP^2) |_P$. We write $L_{st}(X) = T(X) R(X)$ for $R(X) \in K[X]$. Then $| L_{st} |_P = | T |_P \, | R |_P$ and $| R(w) |_P \leqq | R |_P$ ($| w |_P = 1$). Moreover $| L_{st} |_P \leqq 1$ and so $| R |_P \leqq | T |_P^{-1} \leqq \| T \|^{\nu}$. These inequalities may be combined to yield an upper bound for $| L_{st}(w) |_P = | T(w) |_P \, | R(w) |_P$ which would give the result

$$| f^{(s)}(tP^2) |_P < P^{-Nq \log^{\delta} \log q}$$

for all s, t in our desired range.

8) We now use these results to give an upper bound for $| A_{ij}(w) |_P$. By Lemma 8 we write

$$A_{ij}(w) = \sum_{k=0}^{m-1} b_k f^{(k)}(0)$$

where $m = (N+1)(q+1)$. Then by Lemma 9

$$| b_k |_P \leqq P^{c_5 m}.$$

Also, by Theorems 1, 8 of the appendix

$$|f^{(s)}(0)|_P = |s! \int_{0,P} \frac{f(z)z}{z^{s+1}}\, dz\,|_P \leqq P^s \max_{|z|_P = P^{-1}} |f(z)|_P.$$

So fix z such that $|z|_P = 1/P$. Choose $\rho \in \boldsymbol{T}_P$ such that $1/P^2\bar{r} > |\rho|_P > 1/2P^2\bar{r}$ and then we see as an application of Theorems 7, 13 and the corollary to Theorem 4 of the appendix

$$\begin{aligned} f(z) = \int_{0,1} & \left[\frac{z(z-P^2)\cdots(z-\bar{r}P^2)}{\xi(\xi-P^2)\cdots(\xi-P^2)}\right]^{\bar{r}_1+1} \frac{f(\xi)\xi}{\xi - z}\, d\xi \\ & -\sum_{k=0}^{\bar{r}_1}\sum_{t=0}^{\bar{r}} \frac{f^{(k)}(tP^2)}{k!} \int_{tP^2,\rho} \left[\frac{z(z-P^2)\cdots(z-\bar{r}P^2)}{\xi(\xi-P^2)\;\cdots(\xi-\bar{r}P^2)}\right]^{\bar{r}_1+1} \\ & \qquad\qquad \times \frac{(\xi-tP^2)^k(\xi-tP^2)}{\xi-z}\, d\xi. \end{aligned}$$

One sees that the first integral is bounded by $P^{-(\bar{r}_1+1)\cdot(\bar{r}+1)}$. Also one can check that any one of the terms in the second sum is bounded by

$$|f^{(s)}(tP^2)|_P \leqq P^{-3(\bar{r}_1+1)(\bar{r}+1)}.$$

Combining all of these estimates we see

$$|A_{ij}(w)|_P \leqq P^{4(C_5-1/\epsilon)Nq} \leqq P^{-Nq}$$

since we may choose ϵ so that $1/\epsilon - C_5 \geqq 1/4$.

9) It now follows exactly as it did in 6) that for all i, j, $T(X)$ divides $A_{ij}(X)$. This contradicts our assumption in 1) that the $A_{ij}(X)$ are relatively prime.

10) In order to obtain the final result we must solve the following equations:

$$n + \log H = \frac{y}{\sqrt{x \log y}} \log^{-\gamma} \log x$$

$$n + \log^{\sigma} \log H = \sqrt{\frac{x}{\log y}} \log^{-\gamma} \log x,$$

where γ and σ are assumed to be small and x, y are assumed to be large. These may be solved to yield

$$x = (n + \log^{\sigma} \log H)^2 \log^{1+\gamma_1}(n + \log H)$$

$$y = (n + \log^{\sigma} \log H)(n + \log H) \log^{1+\gamma_2}(n + \log H)$$

where $\gamma_1, \gamma_2 \to 0$ as $\gamma \to 0$.

11) We let $H = | G |$ and $n = \deg G$, $[x] = q$ and $[y] = N$. Then it can be checked that $N \geqq q \geqq \log^{1+\epsilon} N$ for some $\epsilon > 0$ sufficiently small. From the equations for x and y and the fact that $\log x \leqq n + \log H$ we see there exists an $\epsilon > 0$ such that

$$yx \log^{\gamma} \log x = (n + \log H)(n + \log^{\sigma} \log H)^3 \log^{2+\epsilon}(n + \log H).$$

Since all three of the conditions of 4 can not hold and the last two do, we see

$$| G(w) |_P > P^{-n^3(n+\log H)\log^{2+\epsilon}(n+\log H)}$$

for all $\epsilon > 0$ sufficiently small provided $n \geqq \log^{\sigma} \log H$. If $n < \log^{\sigma} \log H$ and $n + \log H$ is larger than some constant, we see

$$(n + \log^{\sigma} \log H) \leqq \log^{\gamma_3}(n + \log H)$$

making the above result valid in this case also.

12) The assumption that G is irreducible may be removed by using Lemma 7 of Sec. 5.

This completes the proof.

Appendix. *P*-adic analytic functions. We now give a brief resumé of P-adic analytic functions. For more details see Loonstra [10], Mahler [12], or Bachman [17].

First we recall that a series $\sum a_n$ in $\boldsymbol{T}_P$ converges if and only if $\lim | a_n |_P = 0$ $(n \to \infty)$. We define a P-adic analtic function $f(z)$ as a power series

$$f(z) = \sum_{n=0}^{\infty} a_n (z - a)^n$$

where $a, a_n \in \boldsymbol{T}_P$ for all n, which has a positive radius of convergence about a. It is viewed as a $\boldsymbol{T}_P$-valued function on its domain of convergence. Then f being the uniform limit of continuous functions is continuous. Also if f converges at $x_0 \in \boldsymbol{T}_P$ then f converges at all $x \in \boldsymbol{T}_P$ such that $| x - a |_P \leqq | x_0 - a |_P$ and thus f converges in some open and closed disc about a. We define the derivative of f formally as usual and note $f'(z)$ has precisely the same domain of convergence as does f.

We define a quotient $f(z) = g(z)/G(z)$, where $g(z)$ is an analytic function and $G(z)$ is a polynomial, to be a meromorphic function. Its derivative and the notion of zeros and poles are defined as usual.

We have Taylor's theorem; that is, we may expand an analytic function about any other point in its domain of convergence with the usual formula.

If f is a meromorphic function we may obtain its Laurent expansion about any pole. We define the residue of f at the pole in the usual way. We may add and multiply analytic functions.

It is convenient to separate out a special class of analytic functions (following Mahler).

Definition. We will call $f(z)=\sum_{n=0}^{\infty}a_n(z-a)^n(|a|_P\leqq 1)$ a *normal function* if and only if

1) $|a_n|_P\leqq 1$ for all n
2) $\lim_{n\to\infty}|a_n|_P=0$.

Then given any z such that $|z-a|_P\leqq 1$ (if and only if $|z|_P\leqq 1$) we see $f(z)$ exists and $|f(z)|_P\leqq 1$. Also given any analytic function $g(z)$ we may find integers r, s such that $P^r g(P^s z)$ is a normal function (in fact any r, s sufficiently large work). Now suppose $b\in \boldsymbol{T}_P$ and $|b|_P\leqq 1$. Then

$$f(z)=\sum_{n=0}^{\infty}\frac{f^{(n)}(b)}{n!}(z-b)^n$$

and we can check that the series above is also normal. Also if $f(z)$ is normal, so is $f'(z)$.

Schnirelman integrals. We now develop an analogue of the complex line integral for $\boldsymbol{T}_P$ which will give us the standard results from complex analysis, e.g., Cauchy Integral Formula, Residue Theorem, and Maximum Modulus Principle. It was developed by Schnirelman to prove a P-adic analogue to the Weierstrass Factorization Theorem. The exposition given here is a modification of that given in Loonstra [10]. We note that instead of $\boldsymbol{T}_P$ we could consider any non-archimedian, complete, algebraically closed field.

For all integers n such that $|n|_P=1$ set

$$g_n(X)=X^n-1=(X-\xi_1^{(n)})(X-\xi_2^{(n)})\cdots(X-\xi_n^{(n)}).$$

Clearly for all i,n $|\xi_i^{(n)}|_P=1$. For now, fix $a,r\in \boldsymbol{T}_P$ and consider for all n as above the set $a+r\xi_1^{(n)},\cdots,a+r\xi_n^{(n)}$. Such a sequence of points in $\boldsymbol{T}_P$ is called a *discrete circle with center a and radius $|r|_P=\rho$.*

Definition. Schnirelman Integral. Let $f(x)$ be a $\boldsymbol{T}_P$-valued function defined for all points on the above discrete circle. Then define, if it exists,

$$\int_{a,r} f = \int_{a,r} f(x)\,dx = \lim_{\substack{n\to\infty \\ |n|_P=1}} 1/n \sum_{k=1}^{n} f(a + r\xi_k^{(n)}).$$

(Note: from now on we will write lim for $\lim\limits_{\substack{n\to\infty \\ |n|_P=1}}$.)

We will first note that the usual method of estimating integrals is valid in this context.

THEOREM 1. *Let* $a, r \in \boldsymbol{T}_P$ *and* $f(z)$ *be a function such that*

a) *for all* z *such that* $|z|_P = |r|_P$, $f(a+z)$ *is defined*

b) $\int_{a,r} f$ *exists.*

Then

$$\left|\int_{a,r} f(z)\,dz\right|_P \leqq \max_{\substack{n\geqq 0 \\ 1\leqq i\leqq n}} |f(a + r\xi_i^{(n)})|_P \leqq \max_{|z|_P=|r|_P} |f(a+z)|_P.$$

Proof. Trivial.

THEOREM 2. *Let* $f(z) = \sum_{t=0}^{\infty} a_t f_t(z)$ *where the series on the right converges uniformly to* $f(z)$ *for all points* z *such that* $|z-a|_P = |r|_P$. *Suppose that for all* t, $\int_{a,r} f_t$ *exists. Then* $\int_{a,r} f$ *exists and*

$$\int_{a,r} f = \sum_{t=0}^{\infty} a_t \int_{a,r} f_t.$$

(*In particular, the latter sum converges.*)

Proof. The standard proof for this type of theorem works.

LEMMA 3. *Let* p *be any integer such that* $0 < |p| < n$. *Then*

$$\sum_{i=1}^{n} \xi_i^{(n)p} = 0.$$

Proof. The $\xi_i^{(n)}$ form a group and hence the set $\{\xi_1^{(n)}, \cdots, \xi_n^{(n)}\}$ and the set $\{\xi_1^{(n)-1}, \cdots, \xi_n^{(n)-1}\}$ coincide. So it suffices to assume $p > 0$.

Now with $\sigma_1 = x_1 + \cdots + x_n, \cdots, \sigma_n = x_1 \cdots x_n$ the elementary symmetric functions, we have $x_1^p + \cdots + x_n^p$ is a polynomial in $\sigma_1, \cdots, \sigma_p$ with no constant term (see van der Waerden [15]). Thus since $p < n$ and $\sigma_j(\xi_1^{(n)}, \cdots, \xi_n^{(n)}) = 0$ for $1 \leqq j \leqq n-1$ we have the desired result.

Now let $f(z) = a_0 + a_1 z + a_2 z^2 + \cdots$ be a power series converging for all z such that $|z|_P < R$ $(R > 0)$.

THEOREM 4. *If* $|a|_P < R$ *and* $|r|_P < R$ *then*

$$\int_{a,r} f = f(a)$$

and exists.

Proof. By Theorem 2 it suffices to show that $\int_{a,r} x^k dx = a^k$ for all $k \geqq 0$. Then

$$\begin{aligned}\int_{a,r} x^k dx &= \lim 1/n \sum_{i=1}^{n} (a + r\xi_i^{(n)})^k \\ &= \lim 1/n \sum_{i=1}^{n} \sum_{j=0}^{k} \binom{k}{j} a^{k-j} r^j \xi_i^{(n)j} \\ &= \lim 1/n \sum_{j=0}^{k} \binom{k}{j} a^{k-j} r^j \sum_{i=1}^{n} \xi_i^{(n)j} = a^k\end{aligned}$$

for all $n > k$ by Lemma 3.

COROLLARY. *With the same hypotheses as the theorem*

$$\int_{a,r} f(z)\,(z - a)\,dz = 0.$$

This then would be the analogue of the Cauchy Integral Theorem.

LEMMA 6.
$$\int_{0,r} \frac{dz}{(z-x)^p} = \begin{cases} 0, \ \textit{for} \ |x|_P < |r|_P \\ \dfrac{(-1)^p}{x^p} \ \textit{for} \ |x|_P > |r|_P. \end{cases}$$

Proof. $1/(z-x)^p$ is an analytic function for all z such that $|z|_P < |x|_P$. Thus this case follows from Theorem 4. If $y \in \boldsymbol{T}_P$ and $|y|_P < 1$ we note that

$$\frac{1}{(1-y)^p} = \sum_{t=0}^{\infty} a_t y^t$$

where the a_t are integers. This follows because the binomial theorem, of which this is a special case, is a formal identity in power series and the above series converges in $\boldsymbol{T}_P$ (we note that the a_t are certain binomial coefficients).

Thus if $|x|_P < |z|_P$ we have

$$\frac{1}{(z-x)^p} = \frac{1}{z^p} \frac{1}{(1 - x/z)^p} = \sum_{t=0}^{\infty} a_t \frac{x^t}{z^{p+t}}.$$

So we need only compute (by Theorem 2) for $k \geqq 1$

$$\int_{0,r} \frac{dz}{z^k} = \lim 1/n \sum_{i=0}^{n} \frac{1}{(r\xi_i^{(n)})^k} = 0$$

for all $n > k$ by Lemma 3.

THEOREM 7 (*Analogue of the Cauchy Integral Formula*). *Let* $f(z) = a_0 + a_1 z + \cdots$ *be a power series converging for all* z *such that* $|z|_P < R$ $(R > 0)$. *Suppose* $x, a, r \in T_P$ *are such that* $|x|_P$, $|a|_P$ *and* $|r|_P < R$. *Then*

$$\int_{a,r} \frac{f(z)(z-a)\,dz}{z-x} = \begin{cases} f(x) & \text{for } |x-a|_P < |r|_P \\ 0 & \text{for } |x-a|_P > |r|_P. \end{cases}$$

Proof. Sending $z \to z + a$ we may assume that $a = 0$. Then

$$\int_{0,r} \frac{f(z)z\,dz}{z-x} = \int_{0,r} \left(\sum_{k=0}^{\infty} \frac{a_k z^{k+1}}{z-x} \right) dz = \sum_{k=0}^{\infty} \int_{0,r} \frac{a_k z^{k+1}dz}{z-x}.$$

Also

$$\begin{aligned} \int_{0,r} \frac{z^{k+1}dz}{z-x} &= \int_{0,r} \frac{z^{k+1} - x^{k+1} + x^{k+1}}{z-x}\,dz \\ &= \int_{0,r} (z^k + xz^{k-1} + \cdots + x^k)\,dz + \int_{0,r} \frac{x^{k+1}}{z-x}\,dz \\ &= x^k + x^{k+1} \int_{0,r} \frac{dz}{z-x} \qquad \text{(by Theorem 4).} \end{aligned}$$

The result now follows immediately from Lemma 6.

THEOREM 8. *With the same hypothesis as in Theorem 7 we have*

$$f^{(n)}(x) = n! \int_{a,r} \frac{f(z)(z-a)\,dz}{(z-x)^{n+1}} \quad \textit{for } |x-a|_P < |r|_P.$$

Proof. We may again assume $a = 0$. We have

$$f^{(n)}(x) = \sum_{k=0}^{\infty} \frac{(n+k)!}{k!} a_{n+k} x^k$$

and

$$\begin{aligned} n! \int_{0,r} \frac{f(z)z\,dz}{(z-x)^{n+1}} &= n! \int_{0,r} \frac{a_0 z + a_1 z^2 + \cdots}{(z-x)^{n+1}}\,dz \\ &= \sum_{k=0}^{\infty} n! \int_{0,r} \frac{a_k z^{k+1}dz}{(z-x)^{n+1}}. \end{aligned}$$

Case I. $k < n$.

$$\int_{0,r}\frac{z^{k+1}dz}{(z-x)^{n+1}}=\int_{0,r}\frac{A_1}{z-x}\,dz+\cdots+\int_{0,r}\frac{A_{n+1}}{(z-x)^{n+1}}\,dz$$

where the A_i are constants depending only on x. By Lemma 6 the integrals on the right are all zero.

Case II. $k\geqq n$, say $k=n+h$ and that $h\geqq 0$.

$$\int_{0,r}\frac{z^{n+h+1}}{(z-x)^{n+1}}\,dz$$
$$=\int_{0,r}[z^h+(n+1)z^{h-1}x+\cdots+(n+1)\cdots(n+h)x^h]\,dz$$
$$+\int_{0,r}\frac{G(z)}{(z-x)^{n+1}}\,dz$$

where G is a polynomial of degree $<n+1$. The latter integral is zero, again by the method employed in Case I, and by Theorem 4 the former is $(n+1)\cdots(n+h)x^h$.

Combining Cases I and II gives the result.

THEOREM 9 (Maximum Modulus Principle). *Let* $a, r\in T_P$ $(r\neq 0)$ *and let* $f(z)$ *be a function defined for all* z *such that* $|z-a|_P\leqq |r|_P$. *Suppose that for all* x *such that* $|x-a|_P<|r|_P$ *we have*

$$\int_{a,r}\frac{f(z)\,(z-a)}{z-x}\,dz=f(x).$$

Suppose that for all x *on the discrete circle determined by* a *and* r *we have*

$$|f(x)|_P<M.$$

Then for all z *such that* $|z-a|_P<|r|_P$ *we have*

$$|f(z)|_P<M.$$

Moreover if f *is analytic in the above region we have for all* n

$$\left|\frac{f^{(n)}(z)}{n!}\right|_P<\frac{M}{|r|_P{}^n}.$$

Proof. Trivial.

THEOREM 10 (Liouville Theorem). *An entire analytic function* $f(z)$ *(i.e.,* $f(z)=a_0+a_1z+\cdots$, *the series converging for all* $z\in T_P$*) which is bounded on all of* T_P *is a constant function.*

Proof. Use Theorem 9.

We will now work toward proving a residue theorem.

LEMMA 11. *Let* $F(z)$ *and* $G(z)=(z-a_1)^{\nu_1}\cdots(z-a_k)^{\nu_k}$ *be polynomials such that* $G(z)$ *has no root of absolute value* $|r|_P$ *where* $r\in\mathbf{T}_P$. *Then* $\int_{0,r}\frac{F(z)z}{G(z)}dz=$ *sum of the residues of* $\frac{F(z)}{G(z)}$ *over all poles* z *such that* $|z|_P<|r|_P$.

Proof. We may write

$$\frac{F(z)}{G(z)}=F_1(z)+\frac{A_{11}}{z-a_1}+\cdots+\frac{A_{1\nu_1}}{(z-a_1)^{\nu_1}}+\cdots+\frac{A_{k1}}{z-a_k}+\cdots+\frac{A_{k\nu_k}}{(z-a_k)^{\nu_k}}.$$

Then if say $a_1,\cdots,a_h$ are the roots of G that have absolute value $<|r|_P$ we have

$$\int_{0,r}\frac{F(z)z}{G(z)}dz=A_{11}+\cdots+A_{h1}$$

as desired.

LEMMA 12. *With the same setting as Lemma* 11, *assume also we take* $t\in\mathbf{T}_P$ *such that for* $i=1,2,\cdots,h$ *the interior of the circle of radius* $|t|_P$ *and center* a_i *contains no other* a_j *for* $1\leqq j\leqq k$ *and* $j\neq i$. *Then*

$$\int_{0,r}\frac{F(z)z}{G(z)}dz=\int_{a_1,t}\frac{F(z)(z-a_1)}{G(z)}dz+\cdots+\int_{a_h,t}\frac{F(z)(z-a_h)}{G(z)}dz.$$

Proof. It is obvious that

$$\int_{a_i,t}\frac{F(z)(z-a_i)}{G(z)}dz=A_{i1}\text{ for }1\leqq i\leqq h.$$

Thus the lemma is proved.

THEOREM 13. *Let* $f(z)=c_0+c_1z+\cdots$ *be a power series convergent for all* z *such that* $|z|_P<R$ $(R>0)$. *Let* $r\in\mathbf{T}_P$ *such that* $|r|_P<R$. *Let* $G(z)=(z-a_1)^{k_1}\cdots(z-a_n)^{k_n}$ *be a polynomial such that* $|a_i|_P<|r|_P$ *for* $1\leqq i\leqq n$. *Let* $t\in\mathbf{T}_P$ $(t\neq 0)$ *such that for all* i,j $(i\neq j)$ $|a_i-a_j|_P>|t|_P$. *Then*

$$\int_{0,r}\frac{f(z)z}{G(z)}dz=\int_{a_1,t}\frac{f(z)(z-a_1)}{G(z)}dz+\cdots+\int_{a_n,t}\frac{f(z)(z-a_n)}{G(z)}dz.$$

Proof. Write

$$f(z) = f_1(z) + (z - a_1)^{k_1}\bar{f}(z)$$

where

$$f_1(z) = b_0 + b_1(z - a_1) + \cdots + b_{k_1-1}(z - a_1)^{k_1-1}$$

and

$$\bar{f}(z) = b_{k_1} + b_{k_1+1}(z - a_1) + \cdots$$

where we rewrite

$$f(z) = b_0 + b_1(z - a_1) + \cdots.$$

The proof will be by induction on n.

For $n = 1$ we want

$$\int_{0,r} \frac{f(z)z}{(z - a_1)^{k_1}}\,dz = \int_{a_1,t} \frac{f(z)(z - a_1)}{(z - a_1)^{k_1}}\,dz$$

and by Theorem 8 they are both equal to

$$\frac{1}{(k_1 - 1)!} f^{(k_1-1)}(a_1).$$

For $n > 1$ we have

$$\int_{0,r} \frac{f(z)z}{G(z)}\,dz = \int_{0,r} \frac{f_1(z)z}{G(z)}\,dz + \int_{0,r} \frac{\bar{f}(z)z}{(z - a_2)^{k_2} \cdots (z - a_n)^{k_n}}\,dz$$

$$= \int_{0,r} \frac{f_1(z)z}{G(z)}\,dz + \sum_{i=2}^{n} \int_{a_i,t} \frac{\bar{f}(z)(z - a_i)}{(z - a_2)^{k_2} \cdots (z - a_n)^{k_n}}\,dz$$

(by induction)

$$= \int_{0,r} \frac{f_1(z)z}{G(z)}\,dz + \sum_{i=2}^{n} \int_{a_i,t} \frac{f(z)(z - a_i)}{G(z)}\,dz - \sum_{i=2}^{n} \int_{a_i,t} \frac{f_1(z)(z - a_i)}{G(z)}\,dz$$

$$= \sum_{i=1}^{n} \int_{a_i,t} \frac{f_1(z)(z - a_i)}{G(z)}\,dz + \sum_{i=2}^{n} \int_{a_i,t} \frac{f(z)(z - a_i)}{G(z)}\,dz - \sum_{i=2}^{n} \int_{a_i,t} \frac{f_1(z)(z - a_i)}{G(z)}\,dz$$

(Lemma 12)

$$= \sum_{i=2}^{n} \int_{a_i,t} \frac{f(z)(z - a_i)}{G(z)}\,dz + \int_{a_1,t} \frac{f_1(z)(z - a_1)}{G(z)}\,dz.$$

But

$$\int_{a_1,t} \frac{\bar{f}(z)(z-a_1)^{k_1}(z-a_1)}{G(z)}\,dz = 0$$

by the corollary to Theorem 4. The result follows.

COROLLARY. *Let* $g(z) = f(z)/G(z)$ *where* f *and* G *are as before. Let* $a, r \in \boldsymbol{T}_P$ *such that* $|a|_P$ *and* $|r|_P < R$. *Then*

$$\int_{a,r} g(z)(z-a)\,dz = \sum_{\substack{z \text{ ranges over} \\ \text{all poles } z \text{ such} \\ \text{that } |z-a|_P < |r|_P}} \operatorname{res} g(z).$$

Proof.

$$\int_{a,r} g(z)(z-a)\,dz = \int_{0,r} g(z+a)z\,dz = \sum_{i=1}^{h} \int_{a_i-a,t} g(z+a)(z-a_i+a)\,dz$$

$$= \sum_{i=1}^{h} \operatorname{res}_{a_i-a} g(z+a) = \sum_{i=1}^{h} \operatorname{res}_{a_i} g(z).$$

As one application of the Schnirelman integral a theorem will be proved which was used extensively in the previous sections. It corresponds, in the complex case, to the standard inequality for the absolute value of an analytic function with a large number of zeros of high order.

THEOREM 14 (Mahler). *Let* $f(z)$ *be a normal function vanishing at* $x_1, \cdots, x_n$ *in* $\boldsymbol{T}_P$ *of orders* $d_1, \cdots, d_n$ *respectively. Assume* $|x_i|_P \leqq P^{-r}$ *where* r *is a positive integer and let* $d = d_1 + \cdots + d_n$. *Then given any* $x \in \boldsymbol{T}_P$ *such that* $|x|_P \leqq P^{-r}$ *and any* $k = 0, 1, 2, \cdots$ *we have*

$$|f^{(k)}(x)|_P \leqq P^{-r(d-k)}.$$

Proof. We have by Theorem 8 that

$$f^{(k)}(x) = k! \int_{x,P^r} \frac{f(z)(z-x)}{(z-x)^{k+1}}\,dz$$

so by Theorem 1

$$|f^{(k)}(x)|_P \leqq \max_{|z|_P = P^{-r}} \left|\frac{f(z+x)}{z^k}\right|_P \leqq P^{rk} \max_{|z|_P = P^{-r}} |f(z+x)|_P.$$

Now we noted that if f expanded about one point of absolute value $\leqq 1$ is a normal function then expanding f about any other point of absolute value $\leqq 1$ we still have a normal function. Thus we may write

$$f(z) = (z-x_1)^{d_1} \cdots (z-x_n)^{d_n} g(z)$$

where g is also a normal function. Whence

$$\begin{aligned}|f(z+x)|_P &= |z+x-x_1|_P^{d_1}\cdots|z+x-x_n|_P^{d_n}|g(z+x)|_P\\ &\leqq |z+x-x_1|_P^{d_1}\cdots|z+x-x_n|_P^{d_n}\\ &\leqq P^{-rd}\end{aligned}$$

for all $z \in \boldsymbol{T}_P$ such that $|z|_P = P^{-r}$. Thus

$$|f^{(k)}(x)|_P \leqq P^{rk} \max_{|z|_P = P^{-r}} |f(z+x)|_P \leqq P^{-rd+rk}$$

as desired.

THEOREM 15. *Suppose* $f(z) = \sum_{n=0}^{\infty} a_n(z-a)^n$ *is an analytic function with radius of convergence* $R > 0$ *(*$R = \infty$ *is permissible). Let* b *be in* $\boldsymbol{T}_P$ *such that* $|b-a|_P < R$ *and expand* $f(z)$ *about* b:

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(b)}{n!}(z-b)^n.$$

Then the radius of convergence of this second series is also R. *So the set of convergence of the two series is also the same.*

Proof. It suffices to show that if $|z-b|_P < R$ then the second series converges. For assuming this, then if $|z-b|_P > R$ and the second series converge at z we would have by symmetry $\sum a_n(z-a)^n$ would converge also contradicting the fact that R is the radius of convergence of $\sum a_n(z-a)^n$.

So suppose $|z-b|_P < R$. Then we wish to show

$$\lim_{n\to\infty} \left|\frac{f^{(n)}(b)}{n!}(z-b)^n\right|_P = 0.$$

Choose r in $\boldsymbol{T}_P$ such that $|z-b|_P$ and $|b-a|_P < |r|_P < R$. Then by Theorem 9 we have

$$\left|\frac{f^{(n)}(b)}{n!}\right|_P \leqq \frac{M}{|r|_P^n}$$

where $M = \max_{|x|_P = |r|_P} |f(x)|_P$. Thus

$$\left|\frac{f^{(n)}(b)}{n!}(z-b)^n\right|_P \leqq M\left|\frac{z-b}{r}\right|_P^n$$

and since $\left|\frac{z-b}{r}\right|_P < 1$ the result follows.

Finally as in complex analysis we have the following theorem.

THEOREM 16. *Let* f *be a function meromorphic on some disc* $\mathfrak{D}$. *Let* $z_1, z_2, \cdots$ *be a sequence of distinct points in* $\mathfrak{D}$ *converging to* z_0 *in* $\mathfrak{D}$.

Suppose that all the z_i, $i = 0, 1, 2, \cdots$ *are not among the poles of* f. *Then if* $f(z_i) = 0$ *for all* $i = 1, 2, \cdots$, *we have* $f(z) \equiv 0$.

Proof. Write $f(z) = g(z)/G(z)$ where $g(z)$ is an analytic function and $G(z)$ is a polynomial. Also write

$$g(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n,$$

which converges in some neighborhood of z_0. Since z_0 is not a pole of f we see that $g(z_i) = 0$ for all $i = 1, 2, \cdots$. Exactly as in complex variables we see $a_n = 0$ for all $n = 0, 1, 2, \cdots$ (g is continuous). Thus g vanishes identically in some neighborhood of z_0. But then g must vanish identically in its entire domain of convergence since we have by Theorem 15 that the set of all z such that $\sum a_n (z - z_0)^n$ converges is the entire domain of convergence of g. This completes the proof.

The logarithmic and exponential functions. A brief survey of the logarithmic and exponential functions will be given here (see Artin and Tate [1], pp. 153-156).

A brief survey of the logarithmic and exponential functions will be given here (see Artin and Tate [1], pp. 153-156).

Definition.

$$\log(1 + z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \cdots$$

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots.$$

The properties of these functions are listed below.

1) $\log(1 + z)$ converges iff $|z|_P < 1$.
2) $|z|_P < P^{-1/(P-1)} \rightarrow |\log(1 + z)|_P = |z|_P$.
3) $|a - 1|_P < 1$ and $|b - 1|_P < 1 \rightarrow \log ab = \log a + \log b$.
4) $\log a = 0$ iff a is a P^s root of unity for some s.
5) e^z converges iff $|z|_P < P^{-1/(P-1)}$.
6) $|z|_P < P^{-1/(P-1)} \rightarrow |e^z - 1|_P = |z|_P$.
7) $|a|_P < P^{-1/(P-1)}$ and $|b|_P < P^{-1/(P-1)} \rightarrow e^a e^b = e^{a+b}$.
8) $|z|_P < P^{-1/(P-1)} \rightarrow e^{\log(1+z)} = 1 + z$ and $\log e^z = z$.
9) Let G be the additive subgroup of all $z \in \boldsymbol{T}_P$ such that $|z|_P < P^{-1/(P-1)}$. Then $1 + G$ is a multiplicative subgroup of $\boldsymbol{T}_P{}^*$.

We have a metric on G and $1+G$. Then we see $e^z\colon G \to 1+G$ and $\log\colon 1+G \to G$ are inverse isometric isomorphisms $G \cong 1+G$.

10) For the values of $\log(1+z)$ where $P^{-1/(P-1)} \leqq |z|_P < 1$, we see that log is unbounded. In fact one may show that

$$\lim_{\substack{n\to\infty \\ |n|_P=1}} |\log(1+P^{1/n})|_P = \infty,$$

where $P^{1/n}$ is any n-th root of P.

COLUMBIA UNIVERSITY.

REFERENCES.

[1] E. Artin and J. Tate, *Class Field Theory*, Harvard University, 1954.

[2] A. Gelfond, *Algebraic and transcendental Numbers*, Dover, New York, 1960.

[3] A. Gunther, "Über Transzendente P-adische Zahlen I," *Journal für die Reine und die Angewandte Mathematik*, vol. 192 (1953), pp. 155-166.

[4] ———, "Über Transzendente P-adische Zahlen II," *Journal für die Reine und die Angewandte Mathematik*, vol. 193 (1954), pp. 1-10.

[5] O. Içen, "Eine Verallgemeinerung und Übertragung der Schnerderschen Algebraizitätskriterun ins P-adische mit Anwendung auf einen Transzendenzbeweis im P-adische," *Journal für die Reine und die Angewandte Mathematik*, vol. 198 (1957), pp. 28-55.

[6] ———, "Über die Funktionswerte der P-adischen Reihe

$$\sum_{\nu=0}^{\infty} P a\binom{\nu}{2} z^{\nu}.$$

Journal für die Reine und die Angewandte Mathematik, vol. 202 (1959), pp. 100-106.

[7] S. Lang, *Algebraic Number Theory*, Addison-Wesley, Reading, Massachusetts, 1964.

[8] ———, "Transcendental points on group varieties," *Topology*, vol. 1 (1962), pp. 313-318.

[9] W. LeVeque, *Topics in Number Theory*, Addison-Wesley, Reading, Massachusetts, 1956.

[10] F. Loonstra, *Analytische Untersuchungen über Bewerte Körper*, Amsterdam, 1941.

[11] E. Lutz, *Sur les Approximations Diophantiennes Linéaires P-adiques*, Hermann, Paris, 1955.

[12] K. Mahler, "Über Transzendente P-adische Zahlen," *Compositio Mathematica*, vol. 2 (1935), pp. 259-275.

[13] ———, "On some inequalities for polynomials in several variables," *Journal of the London Mathematical Society*, vol. 37 (1962), pp. 341-344.

[14] C. Siegel, "Transcendental numbers," *Annals of Mathematics Studies*, vol. 16 (1949), Princeton.

[15] B. van der Waerden, *Modern Algebra*, Ungar, New York, 1953.

[16] G. Veldkamp, "Ein Transzendenz Satz für *P*-adische Zahlen," *Journal of the London Mathematical Society*, vol. 15 (1940), pp. 183-192.

[17] G. Bachman, *Introduction to p-adic Numbers and Valuation Theory*, Academic Press, New York, 1964.

[18] S. Lang, "Report on diophantine approximations," *Bulletin de la Société Mathématique de France*, vol. 93 (1965), pp. 177-192.

[19] W. Adams, *Transcendental Numbers in the P-adic Domain*, thesis, Columbia University, 1964.

PAIRS OF 3-CELLS WITH INTERSECTING BOUNDARIES IN E^3.

By C. E. BURGESS.*

Let S_1 and S_2 be 2-spheres in E^3 such that $S_2 + \operatorname{Int} S_2$ is a 3-cell, $S_1 \subset S_2 + \operatorname{Int} S_2$, and S_1 is locally tame from $\operatorname{Int} S_1$ at each point of $S_1 - S_2$. It is the main purpose of this paper to adapt the methods of [4] to show that $S_1 + \operatorname{Int} S_1$ is a 3-cell. A slightly more general result is given in Theorem 1. Investigation of this problem resulted from a question raised by B. G. Casler during the discussion when [4] was presented at the Summer Meeting of the American Mathematical Society in Boulder in August, 1963. His question is answered specifically with the corollary following Theorem 2. In Theorem 4, these results are generalized where S_1 and S_2 are replaced with connected 2-manifolds in a connected 3-manifold. These methods and results enable us to give conditions under which a disk in a 3-cell is locally tame from one side (Theorem 3).

We will use notation and definitions similar to those in [4]. If S is a 2-sphere in E^3 and U is a component of $E^3 - S$, we define S to be *locally tame from U at the point p* of S if there is a subset V of $S + U$ such that $p \in V$, V is open relative to $S + U$, and the closure of V is a 3-cell. A similar definition of "locally tame from a complementary domain" can be given for a connected 2-manifold which separates a connected 3-manifold. If D is a disk in E^3, we define $\operatorname{Int} D$ to be *locally tame from one side* if, for each point p of $\operatorname{Int} D$, there exists a disk D' and a 3-cell K such that $p \in \operatorname{Int} D' \subset D$ and $K \cdot D = D' \subset \operatorname{Bd} K$.

THEOREM 1. *Suppose S_1 and S_2 are 2-spheres in E^3 and R_1 and R_2 are disks on S_1 and S_2, respectively, such that*

(1) *S_2 is locally tame from $\operatorname{Int} S_2$ at each point of $\operatorname{Int} R_2$,*
(2) *$R_1 \cdot R_2 \subset \operatorname{Int} R_2$,*
(3) *S_1 is locally tame from $\operatorname{Int} S_1$ at each point of $\operatorname{Int} R_1 - R_2$,*
(4) *$R_2 \subset S_1 + \operatorname{Ext} S_1$, and*
(5) *$R_1 \subset S_2 + \operatorname{Int} S_2$.*

Then S_1 is locally tame from $\operatorname{Int} S_1$ at each point of $\operatorname{Int} R_1$.

Received July 20, 1964.
Revised October 20, 1965.
* This work was supported by the National Science Foundation under GP-107.

Proof. We will show that for each point p of $\operatorname{Int} R_2 \cdot \operatorname{Int} R_1$ and for each neighborhood N of p, there is a neighborhood Q of p such that each unknotted simple closed curve in $Q \cdot \operatorname{Int} S_1$ is the boundary of a disk in $N \cdot \operatorname{Int} S_1$. Then from Theorem 2 of [4] and requirement (3) of the hypothesis, it will follow that $\operatorname{Int} S_1$ is locally simply connected at each point of $\operatorname{Int} R_1$. This implies that S_1 is locally tame from $\operatorname{Int} S_1$ at each point of $\operatorname{Int} R_1$ (See [3] and Theorems 1 and 2 of [5]).

Let X be an unbounded topological ray which has its starting point in $S_1 - R_1$ and which otherwise is a subset of $\operatorname{Ext} S_1$. Let p be a point of $\operatorname{Int} R_1 \cdot \operatorname{Int} R_2$, and let N be a spherical neighborhood of p such that $N \cdot S_1 \subset \operatorname{Int} R_1$, $N \cdot S_2 \subset R_2$, and N does not intersect X. There exists a spherical neighborhood Q of p such that $Q \subset N$, $Q \cdot \operatorname{Int} S_1 \subset \operatorname{Int} S_2$, and the sets $Q \cdot S_1$ and $Q \cdot S_2$ are subsets of the interiors of disks R_1' and R_2' in $N \cdot S_1$ and $N \cdot S_2$, respectively, where $R_1' \cdot S_2 \subset R_2'$. Let J be an unknotted simple closed curve in $Q \cdot \operatorname{Int} S_1$. From the hypothesis that S_1 is locally tame from $\operatorname{Int} S_1$ at each point of $\operatorname{Int} R_1 - R_2$, it follows that $\operatorname{Int} S_1 + \operatorname{Int} R_1 - R_2$ is a 3-manifold with boundary. Hence $\operatorname{Int} S_1 + \operatorname{Int} R_1 - R_2$ can be triangulated, relative to itself [1; 2; 6]. This implies that there is a homomorphism h of S_1 into $S_1 + \operatorname{Int} S_1$ such that

(1) h is the identity on $S_1 \cdot S_2$ and on $S_1 - R_1'$,
(2) $S_1 \cdot h(\operatorname{Int} R_1') = S_2 \cdot h(\operatorname{Int} R_1') = S_2 \cdot \operatorname{Int} R_1'$,
(3) $h(R_1') \subset R_1' + N \cdot \operatorname{Int} S_1$,
(4) $Q \cdot h(S_1) \subset h(R_1')$, and
(5) $J \subset \operatorname{Int} h(S_1)$.

Furthermore, by using Theorem 8 of [2], we can require that

(6) $h(S_1)$ is locally polyhedral in E^3 at each point of $h(S_1) - S_1$.

Since S_2 is locally tame from $\operatorname{Int} S_2$ at each point of R_2 and J is the boundary of a disk in Q, it follows that J is the boundary of a disk D in $N \cdot \operatorname{Int} S_2$. As in the proof of Theorem 3 in [4], we can adjust D so that (1) D is locally polyhedral at each point of $\operatorname{Int} D$, (2) D is in general position with $h(R_1')$, and (3) the number of components of $D \cdot h(R_1')$ cannot be reduced by adjusting D near $D + h(R_1')$. Suppose that, with such an adjustment, D intersects $h(S_1)$. Each component of $D \cdot \operatorname{Ext} h(S_1)$ separates $\operatorname{Ext} h(S_1)$ into two connected sets such that one of them, say E, is a bounded subset of E^3. Since $D \subset N$ and N does not intersect X, it follows that $\operatorname{Bd} E$ does not intersect X. Hence $\operatorname{Bd} E$ is a subset of $D + \operatorname{Int} h(R_1')$. This implies that $\operatorname{cl}(E)$ is a subset of N, and consequently $\operatorname{cl}(E)$ does not intersect $S_2 - R_2$.

This enables us to conclude that S_2 does not intersect $\mathrm{cl}(E)$, as $R_2 \subset h(S_1) + \mathrm{Ext}\, h(S_1)$ and R_2 does not intersect D. Hence we have shown that for each component Y of $D \cdot \mathrm{Ext}\, h(S_1)$, the closure of the bounded component of $\mathrm{Ext}\, h(S_1) - Y$ does not intersect S_2.

Now beginning with the third paragraph of the proof of Theorem 3 in [4], we can follow a similar procedure to obtain a contradiction to the supposition that D intersects $h(S_1)$. Hence $D \subset N \cdot \mathrm{Int}\, h(S_1)$, and since $\mathrm{Int}\, h(S_1) \subset \mathrm{Int}\, S_1$, it follows that J is the boundary of a disk in $N \cdot \mathrm{Int}\, S_1$. As indicated above, this implies that S_1 is locally tame from $\mathrm{Int}\, S_1$ at each point of $\mathrm{Int}\, R_1$.

THEOREM 2. *Suppose S_1 and S_2 are 2-spheres in E^3 such that*

(1) *$S_2 + \mathrm{Int}\, S_2$ is a 3-cell,*
(2) *$S_1 \subset S_2 + \mathrm{Int}\, S_2$, and*
(3) *S_1 is locally tame from $\mathrm{Int}\, S_1$ at each point of $S_1 - S_2$.*

Then $S_1 + \mathrm{Int}\, S_1$ is a 3-cell.

Proof. It follows that from Theorem 1 that S_1 is locally tame from $\mathrm{Int}\, S_1$. This implies that S_1 is tame from $\mathrm{Int}\, S_1$ [1; 6]. Thus $S_1 + \mathrm{Int}\, S_1$ is a 3-cell.

COROLLARY. *Suppose S_1 and S_2 are 2spheres in E^3 such that*

(1) *S_2 is tame,*
(2) *$S_1 \subset S_2 + \mathrm{Int}\, S_2$, and*
(3) *S_1 is locally tame at each point of $S_1 - S_2$.*

Then $S_1 + \mathrm{Int}\, S_1$ is a 3-cell.

THEOREM 3. *If K is a 3-cell in E^3 and D is a disk such that $D \subset K$ and D is locally tame at each point of $D - \mathrm{Bd}\, K$, then $\mathrm{Int}\, D$ is locally tame from one side.*

Proof. Let p be a point of $\mathrm{Int}\, D$. We need to consider only the case where $p \in \mathrm{Bd}\, K$. Let S_2 denote $\mathrm{Bd}\, K$ and let R_2' be a disk on S_2 such that $p \in \mathrm{Int}\, R_2'$ and R_2' does not intersect $\mathrm{Bd}\, D$. Bing [3] has shown that there exist a 2-sphere S_1' and a disk R_1' such that

(1) $p \in \mathrm{Int}\, R_1'$,
(2) $R_1' \subset S_1' \cdot \mathrm{Int}\, D$, *and*
(3) $S_2 \cdot S_1' \subset \mathrm{Int}\, R_2'$.

Now there exists a subdisk R_2'' of R_2' such that $p \in \operatorname{Int} R_2''$ and either $R_2'' \subset S_1' + \operatorname{Int} S_1'$ or $R_2'' \subset S_1' + \operatorname{Ext} S_1'$. In either case we can apply Theorem 1 of [5], which is a slight modification of Bing's theorem cited above, to obtain a 2-sphere S_1 and disks R_1 and R_2 such that

(1) $p \in \operatorname{Int} R_1 \subset R_1'$,
(2) $p \in \operatorname{Int} R_2 \subset R_2''$,
(3) $R_1 \cdot R_2 \subset \operatorname{Int} R_2$,
(4) $R_2 \subset S_1 + \operatorname{Ext} S_1$.

Now it follows readily that the spheres S_1 and S_2 and the disks R_1 and R_2 satisfy all of the requirements in the hypothesis of Theorem 1. Thus S_1 is locally tame from $\operatorname{Int} S_1$ at each point of $\operatorname{Int} R_1$, and this implies that $\operatorname{Int} D$ is locally tame from one side.

THEOREM 4. *Suppose M_1 and M_2 are connected 2-manifolds in a triangulated 3-manifold M^3 such that*

(1) *$M^3 - M_1$ is the sum of two connected sets V_1 and W_1,*
(2) *$M^3 - M_2$ is the sum of two connected sets V_2 and W_2,*
(3) *M_2 is tame from V_2,*
(4) *$M_1 + V_1 \subset M_2 + V_2$, and*
(5) *M_1 is locally tame from V_1 at each point of $M_1 - M_2$.*

Then M_1 is tame from V_1.

Proof. Let p be a point of M_1. We need to consider only the case where $p \in M_2$. Let N be a neighborhood of p that is homeomorphic with E^3. By applying Theorem 1 of [5] as in the proof of Theorem 9 of [4], we can obtain a disk R_2 and a 2-sphere S_2 such that

(1) $S_2 \subset N$,
(2) $p \in \operatorname{Int} R_2 \subset M_2 \cdot S_2$, and
(3) p is on the boundary of some component of $V_2 \cdot \operatorname{Int} S_2$, where $\operatorname{Int} S_2$ denotes the component of $M^3 - S_2$ that is a subset of N.

Requirements (2) and (3) above and requirement (3) in the hypothesis imply that S_2 is locally tame from $\operatorname{Int} S_2$ at each point of $\operatorname{Int} R_2$. With another application of Theorem 1 of [5], we obtain a disk R_1 and a 2-sphere S_1 in N such that

(1) $p \in \operatorname{Int} R_1 \subset S_1 \cdot M_1$,
(2) $S_1 \cdot S_2 \subset R_2$, and
(3) R_1 is on the boundary of some component of $V_1 \cdot \operatorname{Int} S_1$.

Now we can replace each of R_1 and R_2 with a subdisk of itself, without a change in notation, so that the requirements in the hypothesis of Theorem 1 are satisfied. Hence it follows from Theorem 1 that S_1 is locally tame from $\operatorname{Int} S_1$ at each point of $\operatorname{Int} R_1$, and this implies that M_1 is locally tame from V_1. Thus M_1 is tame from V_1 [1;6].

THE UNIVERSITY OF UTAH.

REFERENCES.

[1] R. H. Bing, "Locally tame sets are tame," *Annals of Mathematics*, vol. 59 (1954), pp. 145-158.

[2] ———, "An alternative proof that 3-manifolds can be triangulated," *Annals of Mathematics*, vol. 69 (1959), pp. 37-65.

[3] ———, "A surface is tame if its complement is 1-ULC," *Transactions of the American Mathematical Society*, vol. 101 (1961), pp. 294-305.

[4] C. E. Burgess, "Properties of certain types of wild surfaces in E^3," *American Journal of Mathematics*, vol. 86 (1964), pp. 325-338.

[5] ———, "Characterizations of tame surfaces in E^3," *Transactions of the American Mathematical Society*, vol. 114 (1965), pp. 80-97.

[6] E. E. Moise, "Affine structures in 3-manifolds, VIII; Invariance of the knot types; Local tame imbeddings," *Annals of Mathematics*, vol. 59 (1954), pp. 159-170.

A SPECTRAL THEORY FOR SELF-ADJOINT SINGULAR INTEGRAL OPERATORS.[1]

By Marvin Rosenblum.

1. Introduction. In this paper we shall follow up the work [5], [6], [9], [16], [17] of W. Koppelman and J. D. Pincus on the spectral theory of self-adjoint singular integral operators of Cauchy type. Let R be the real line and E any Lebesgue measurable subset of R of positive or infinite measure. For any p, $1 \leqq p \leqq \infty$, $L^p(E)$ shall be the usual complex Lebesgue space over E. We fix Lebesgue measurable functions a and b on E such that a is bounded below and real and b is in $L^\infty(E)$. We further assume that b does not vanish on any set of positive measure. Our object here is to study the self-adjoint operator S on $L^2(E)$, where

$$Sg(x) = a(x)g(x) + (\pi i)^{-1}b(x)\int_E^* (t-x)^{-1}b^*(t)g(t)\,dt. \tag{1.1}$$

The asterisk by the integral sign indicates that the integral is to be taken as a Cauchy principal value. $b^*(t)$ is the complex conjugate of $b(t)$.

The methods used by Koppelman and Pincus to diagonalize S rather basically involve smoothness conditions on a and b. I shall employ a variation of the theory I propounded for diagonalizing self-adjoint Toeplitz operators in [11], [12], and [13]. We derive an expression for the resolvent operator of S (Theorem 1), and then show that the spectral measure $E_S(\cdot)$ of S is weakly absolutely continuous except for the possibility of eigenvalues when $E = R$ (Theorems 2 and 3). Then S is diagonalized (Theorem 5) and it turns out that the multiplicity $n(\xi)$ of a point ξ in the continuous spectrum is the "index" (defined and studied in Section 4 and Theorem 4) of the set Γ_ξ specified in (5.1). Thus the singular integral operator S behaves spectrally much like the self-adjoint Toeplitz operators of [13] and is amenable to an application of similar techniques.

Finally we consider two bounded self-adjoint operators $\mathfrak{A}$ and $\mathfrak{B}$ in an abstract Hilbert space $\mathfrak{H}$ whose commutator has a one dimensional range, say, all scalar multiples of a vector e in $\mathfrak{H}$, and such that

---

Received January 29, 1965.

[1] This research was supported by National Science Foundation Grant No. G-18853.

$$\{\mathfrak{A}^n e, n = 0, 1, 2, \cdots\}$$

is total in $\mathfrak{H}$. We complement results of Xa Dao-Xeng [4] and provide a complete set of unitary invariants for $\mathcal{S}$ (Theorem 7). The proof depends on the result of [4] that $\mathcal{S}$ is necessarily unitarily equivalent to a singular integral operator S.

I gratefully acknowledge stimulating and helpful discussions and correspondence with Louis de Branges during the course of this work.

After redacting this paper I received a copy of [18] from J. D. Pincus, a work that antedates ours. In [18] Pincus presents a general theory of generalized (distribution-valued) eigenfunction expansions for self-adjoint operators. He then extends the results derived in [17] to diagonalize S under the assumptions that E is a finite interval and a and b are bounded and measurable. His methods differ from ours in that they depend on generalized eigenfunction expansions whereas we use representations for the resolvent operator. In the recent paper [19] C. R. Putnam has results on the spectrum of S when $R\backslash E$ contains an open interval. The theory we present here as the merit of including the difficult case when $E = R$.

2. A generating function formula for the resolvent. We shall be considering S as an operator on $L^2(E)$, but for notational convenience we define the functions a and b on all of R by demanding that they vanish outside of E. We use "$\int$" to mean "$\int_R$." Throughout u and v shall be non real complex numbers.

Since the Hilbert transform has norm one, ([1], p. 120), the operator S is bounded with $\| S \| \leqq \operatorname{ess\,sup} | b |^2$ when $a \equiv 0$. It follows that S is a self-adjoint operator with domain

$$\{f \in L^2(E) : \int | a(t) f(t) |^2 \, dt < \infty\},$$

and S is bounded below. We define the $L^2(E)$ element k_u by

$$k_u(x) = (x - u)^{-1} b(x), \tag{2.1}$$

and similarly define k_v. In this section we shall obtain a simple formula for $\langle (S - \lambda)^{-1} k_u, k_v \rangle$, where $\operatorname{Im} \lambda > 0$. Our method follows Carleman's technique for the solution of integral equations (see [14], p. 186). The formula to be derived is analogous to the formula for the resolvent of a Toeplitz operator derived in [7], p. 238, [3], and [11].

Fix a complex number λ so that $\operatorname{Im} \lambda > 0$ or

$$\lambda < m = \operatorname{ess\,inf}\{a(t) - |b(t)|^2 : t \in E\}$$

and define

$$w(t,\lambda) = (a(t) + |b(t)|^2 - \lambda)(a(t) - |b(t)|^2 - \lambda)^{-1}. \tag{2.2}$$

It is easily checked that, for almost all t, $w(t,\lambda)$ takes values in the upper half plane or in the positive reals. Thus $\log w(t,\lambda)$ is a.e. well-defined, where log is the principal branch of the logarithm function. Furthermore $\log w(\cdot,\lambda)$ is in $L^\infty(R)$ since $w(\cdot,\lambda)$ and $1/w(\cdot,\lambda)$ are. Define for all non-real complex z

$$W(z,\lambda) = \exp\left[\tfrac{1}{2\pi i}\int [(t-z)^{-1} - t(t^2+1)^{-1}] \log w(t,\lambda)\,dt\right] \tag{2.3}$$

$W(\cdot,\lambda)$ enjoys the following properties:

(2a) If $\lambda < m$ then $W(\cdot,\lambda)$ and $1/W(\cdot,\lambda)$ are bounded functions separately holomorphic in the upper and lower half-planes.

(2b) $W(x+\mathrm{i}0,\lambda)/W(x-i0,\lambda) = w(x,\lambda)$ for almost all real x.

THEOREM 1. *Suppose* $\operatorname{Im}\lambda > 0$. *Then*

$$\langle (S-\lambda)^{-1}k_u, k_v\rangle = i\pi(v^* - u)^{-1}[1 - W(u,\lambda)/W(v^*,\lambda)]. \tag{2.4}$$

Proof. S is bounded below, say $S \geqq m_1$. Choose $\lambda < \min(m_1, m)$. Then the $L^2(E)$ integral equation $Sf - \lambda f = k_u$ has a unique solution $f = (S-\lambda)^{-1}k_u$. Necessarily

$$(a(x) - \lambda)b^*(x)f(x) + \frac{1}{\pi i}|b(x)|^2 \int^* (t-x)^{-1} b^*(t)f(t)\,dt \\ = (x-u)^{-1}|b(x)|^2$$

for all real x. Following Carleman we introduce

$$F(z) = \tfrac{1}{2\pi i}\int (t-z)^{-1} b^*(t)f(t)\,dt, \operatorname{Im} z \neq 0.$$

Then

$$F(x+i0) - F(x-i0) = b^*(x)f(x) \text{ and}$$

$$F(x+i0) + F(x-i0) = \frac{1}{\pi i}\int^* (x-t)^{-1} b^*(t)f(t)\,dt.$$

By the use of these equations and (2b) we find that

$$F(x+i0)W(x+i0,\lambda) - F(x-i0)W(x-i0,\lambda) \\ = \tfrac{1}{2}(x-u)^{-1}[W(x+i0,\lambda) - W(x-i0,\lambda)]$$

$$= G(x+i0) - G(x-i0), \text{ where}$$
$$G(z) = \tfrac{1}{2}(z-u)^{-1}[W(z,\lambda) - W(u,\lambda)].$$

Now, $F(\cdot)W(\cdot,\lambda)$ and G belong to the Hardy class H^2 (see [8], p. 38), and we deduce from the above equations (see [15], p. 132) that $F(v^*)W(v^*,\lambda) = G(v^*)$ for all non-real v. Thus

$$\tfrac{1}{2\pi i}\langle(S-\lambda)^{-1}k_u, k_v\rangle = F(v^*)$$
$$= \tfrac{1}{2}(v^*-u)^{-1}[1 - W(u,\lambda)/W(v^*,\lambda)],$$

which implies 2.4 for $\lambda < \min(m_1, m)$. An analytic continuation argument yields (2.4) for $\operatorname{Im}\lambda > 0$.

Corollary. *Let S_1 be the operator S on $L^2(R)$ with a and $|b|^2$ (but not necessarily b) constant functions. Then for all non-real λ,*

$$\langle(S_1-\lambda)^{-1}k_u, k_v\rangle = 2\pi i\epsilon(u,v)|b|^2/[(u-v^*)(a-\epsilon(u,v)|b|^2-\lambda)],$$

where $\epsilon(u,v) = \tfrac{1}{2}[\operatorname{sgn}\operatorname{Im}u + \operatorname{sgn}\operatorname{Im}v]$. Thus the spectrum consists of the points $m = a - |b|^2$ and $M = a + |b|^2$.

Proof. The formula for the resolvent of S_1 follows from 2.4 by direct computation. An analytic continuation of λ in both sides of the formula shows that $\{m, M\}$ are the only spectral singularities of S_1.

3. The absolute continuity of S. In [10] C. R. Putnam proved that in case a is bounded and E is of finite measure, then S is AC on R; that is, the spectral measure E_S of S is weakly absolutely continuous on R. His proof depends on his theory of commutators in Hilbert space and does not appear to be directly applicable to unbounded operators or to the complex situation when $E = R$, where S may have eigenvalues. We shall employ here the Aronszajn-Donoghue theory [2] of exponential representations of the class P of not identically zero functions Φ holomorphic and having non-negative imaginary part in the upper half-plane. Such Φ have the following two representations:

(3a) Let $(\mu[\infty], \beta, \mu)$ be a triple consisting of a non-negative number $\mu[\infty]$, a real number β, and a measure μ on R such that $\int(1+t^2)^{-1}d\mu(t) < \infty$. Further assume that it is not the case that $\mu[\infty]$, β, and μ are all null. Then the representation

$$(3.1)\quad \Phi(\lambda) = \mu[\infty]\lambda + \beta + \int[(t-\lambda)^{-1} - t(t^2+1)^{-1}]d\mu(t), \operatorname{Im}\lambda > 0,$$

provides a one-to-one correspondence between the set of such triples and P.

(3b) Let $(\alpha, \{f\})$ be a pair consisting of a real number and the class of all functions equal a. e. on R to a measurable function f satisfying $0 \leqq f \leqq 1$ a. e. Then the representation

$$(3.2) \qquad \Phi(\lambda) = \exp[\alpha + \int [(\xi - \lambda)^{-1} - \xi(\xi^2 + 1)^{-1}] f(\xi)\, d\xi], \operatorname{Im} \lambda > 0,$$

provides a one-to-one correspondence between the set of such pairs and P. f is determined a. e. by

$$(3.3) \qquad f(\xi) = \frac{1}{\pi} \lim_{\eta \downarrow 0} \operatorname{Im} \log \Phi(\xi + i\eta).$$

From (3a) and (3b) we obtain

(3c) There is a one-to-one correspondence between the triple $(\mu[\infty], \beta, \mu)$ of (3a) and the pair $(\alpha, \{f\})$ of (3b) associated with a given Φ in P.

[2] is devoted to a study of the correspondence of (3c). For our immediate purpose the following corollary of [2], Theorem VIII, p. 386 will prove useful:

(3d) Suppose $\Phi \in P$ and μ and $\{f\}$ are associated with Φ by (3a) and (3b). Suppose that there is an f in $\{f\}$ that is of bounded variation on an open interval (r, s) with $f(\xi + 0) - f(\xi - 0) < 1$ for all ξ in (r, s). Then μ is absolutely continuous on (r, s).

For our application of this result and for later reference it is necessary to consider the limiting values of $W(z, \lambda)$ as $\operatorname{Im} \lambda \downarrow 0$. We set down the following notations:

$$(3.4) \qquad \begin{aligned} \Gamma_\xi &= \{t \in E : -\infty < w(t, \xi) \leqq 0\}, \\ A_\xi &= \{t \in E : a(t) - |b(t)|^2 - \xi \leqq 0\}, \\ {}_0A_\xi &= \{t \in E : a(t) - |b(t)|^2 - \xi = 0\}, \\ B_\xi &= \{t \in E : a(t) + |b(t)|^2 - \xi \leqq 0\}, \text{ and} \\ {}_0B_\xi &= \{t \in E : a(t) + |b(t)|^2 - \xi = 0\}. \end{aligned}$$

Suppose $z = x + iy$ with x, y real and $y \neq 0$. Define the Poisson kernel P_z by $P_z(t) = \pi^{-1} |y| [(t - x)^2 + y^2]^{-1}$.

One can show that the following equations hold for almost all real ξ:

$$(3.5) \qquad \begin{aligned} &\lim_{\eta \downarrow 0} W(z, \xi + i\eta) W^*(z^*, \xi + i\eta) \\ &\qquad = \exp \int_{\Gamma_\xi} [(t - z)^{-1} - t(t^2 + 1)^{-1}]\, dt. \end{aligned}$$

(3.6) If $y > 0$,

$$\frac{1}{\pi}\lim_{\eta\downarrow 0}\operatorname{Im}\log[W(z,\xi+i\eta)/W(z^*,\xi+i\eta)] = \int_{\Gamma_\xi} P_z(t)\,dt = \int_{A_\xi} P_z(t)\,dt - \int_{B_\xi} P_z(t)\,dt.$$

(3.7) If $y < 0$,

$$\frac{1}{\pi}\lim_{\eta\downarrow 0}\operatorname{Im}\log[-W(z,\xi+i\eta)/W(z^*,\xi+i\eta)] = 1 - \int_{\Gamma_\xi} P_z(t)\,dt.$$

It should be noted that $\int_{{}_0A_\xi} P_z(t)\,dt = \int_{{}_0B_\xi} P_z(t)\,dt = 0$ for almost all ξ.

Define Φ_+ and Φ_- on the upper half-plane by

$$\Phi_+(\lambda) = \frac{1}{i\pi}(z - z^*)\langle (S-\lambda)^{-1}k_z, k_z\rangle + 1 \text{ if } y > 0,$$

and

$$\Phi_-(\lambda) = -\frac{1}{i\pi}(z - z^*)\langle (S-\lambda)^{-1}k_z, k_z\rangle - 1 \text{ if } y < 0.$$

By 2.4,

$$(3.8) \qquad \begin{aligned} \Phi_+(\lambda) &= W(z,\lambda)/W(z^*,\lambda), \text{ and} \\ \Phi_-(\lambda) &= -W(z,\lambda)/W(z^*,\lambda). \end{aligned}$$

Φ_+ and Φ_- are holomorphic functions with positive imaginary parts. (3a) and (3b) associate with Φ_+ and Φ_- the measures $\frac{1}{i\pi}(z - z^*)\langle E_S(\cdot)k_z, k_z\rangle$ and $-\frac{1}{i\pi}(z - z^*)\langle E_S(\cdot)k_z, k_z\rangle$ and the functions

$$f_+(\xi) = \int_{A_\xi} P_z(t)\,dt - \int_{B_\xi} P_z(t)\,dt$$
$$f_-(\xi) = 1 - f_+(\xi)$$

respectively. f_+ and f_- are of bounded variation on R. We put

$$m = \operatorname{ess\,inf}\{a(t) - |b(t)|^2 \colon t \in E\} \text{ and}$$
$$M = \operatorname{ess\,sup}\{a(t) - |b(t)|^2 \colon t \in E\}.$$

THEOREM 2. i) *S is AC on R if $R - E$ is not a set of measure zero.*

ii) *Suppose $E = R$. S is AC on R if $a - |b|^2$ and $a + |b|^2$ are not a. e. equal to constant functions.*

iii) *Suppose $E = R$. S is AC on every real interval that does not contain m or M.*

Proof. We apply (3d) to f_+ and f_-.

$$f_+(\xi+0)-f_+(\xi-0)=\int_{{}_0A_\xi}P_z(t)\,dt-\int_{{}_0B_\xi}P_z(t)\,dt.$$

Thus $f_+(\xi+0)-f_+(\xi-0)\leqq 1$ and $f_+(\xi+0)-f_+(\xi-0)=1$ if and only if $E=R$ and $a-|b|^2=m$ a. e. on R. Similarly $f_-(\xi+0)-f_-(\xi-0)\leqq 1$, with equality holding if and only if $E=R$ and $a+|b|^2=M$ a. e. on R.

Suppose now that $\langle E_S(\cdot)k_z,k_z\rangle$ is AC on an interval I for all non-real z. Suppose u and v are non-real complex numbers. The Schwarz inequality shows that

$$|\langle E_S(\Delta)k_u,k_v\rangle|^2\leqq\langle E_S(\Delta)k_u,k_u\rangle\langle E_S(\Delta)k_v,k_v\rangle$$

for all real Borel sets Δ, so $\langle E_S(\cdot)k_u,k_v\rangle$ is AC on I. Since $\{k_u: u \text{ non-real}\}$ is total in $L^2(E)$ it follows that S is AC on I. Thus Theorem 2 follows from the preceding paragraph and (3d).

The gist of Theorem 2 is that S is AC on R except for the possibility of eigenvalues at m or M when $E=R$ and $a-|b|^2=m$ a. e. or $a+|b|^2=M$ a. e. We next explore this possibility, and define closed linear manifolds $\mathfrak{M}_m$ and $\mathfrak{M}_M$ that, when non-trivial, will turn out to be eigenspaces of S. Consider the following conditions:

Ci) $\int(t^2+1)^{-1}\log|b(t)|\,dt>-\infty$.
Cii) $E=R$ and $a-|b|^2=m$ a. e..
Ciii) $E=R$ and $a+|b|^2=M$ a. e..

If Ci) or ii) is false, then $\mathfrak{M}_m=\{0\}$. If Ci) or iii) is false, then $\mathfrak{M}_M=\{0\}$. In the remaining cases Ci) is valid and we define the function B on the upper half-plane by

$$B(z)=\exp[\tfrac{1}{2\pi i}\int[(t-z)^{-1}-t(t^2+1)^{-1}]\log|b(t)|\,dt].$$

If Ci) and ii) hold, then

$$\mathfrak{M}_m=\{(h(x+i0)B^2(x+i0)/b(x))^*: h\in H^2\}.$$

If Ci) and iii) hold, then

$$\mathfrak{M}_M=\{h(x+i0)B^2(x+i0)/b^*(x): h\in H^2\}.$$

THEOREM 3. *Suppose* $E=R$.

i) *S is AC on R if and only if* $\mathfrak{M}_m=\mathfrak{M}_M=\{0\}$.

ii) *Non-null elements of $\mathfrak{M}_m$ and $\mathfrak{M}_M$ are eigenvectors of S corresponding to the eigenvalues m and M respectively.*

Proof. Suppose Cii) and the negation of Ci) hold. By (3.8),

$$\eta\Phi_+(m+i\eta)=\exp\int P_z(t)\log(\eta+2i\,|\,b(t)\,|^2)\,dt,$$

so $\lim\limits_{\eta\downarrow 0}|\,\eta\Phi_-(m+i\eta)\,|=0.$

$$\Phi_-(m+i\eta)=-\eta\exp[-\int P_z(t)\log(\eta+2i\,|\,b(t)\,|^2)\,dt],$$

so $\lim\limits_{\eta\downarrow 0}|\,\eta\Phi_-(m+i\eta)\,|=0.$

Now we shall employ the following result: Let Φ and μ be as in (3a). Then, (as indicated in [2], p. 328, II),

$$\lim_{\eta\downarrow 0} i\eta\Phi(m+i\eta)=\mu(\{m\}).$$

Thus for each non-real z the measure $\langle E_S(\cdot)k_z,k_z\rangle$ has no point mass at m. By the argument concluding the proof of Theorem 2 it follows that $\langle E_S(\cdot)x,y\rangle$ has no point mass at m for each x, y in $L^2(R)$, so m is not an eigenvalue of S.

The situation where Ciii) and the negation of Ci) hold can be similarly treated. In this case M is not an eigenvalue of S.

So far we have shown that S is AC whenever Ci) is false. In the remainder of this proof we assume that Ci) holds, so B is well-defined. B is a bounded holomorphic function in the upper half-plane such that $|\,B(x+i0)\,|^2=|\,b(x)\,|$ a.e.. In addition B enjoys the following property: If an H^2 function F is such that $F(x+i0)/B^2(x+i0)$ is in $L^2(R)$, then F/B^2 is in H^2.

Suppose Ciii) holds. Then an element f in $L^2(R)$ satisfies $Sf=Mf$ if and only if f satisfies each of the following equivalent statements:

i) $b^*(x)f(x)-(\pi i)^{-1}\int^*(t-x)^{-1}b^*(t)f(t)\,dt=0.$

ii) $\int b^*(t)f(t)\,(t-z^*)^{-1}dt=0$ for all complex z with $\operatorname{Im} z>0.$

iii) $b^*f\in H^2.$

Thus any f in $\mathfrak{M}_m$ satisfies $Sf=Mf$. Suppose next that $Sf=Mf$ but $f\notin\mathfrak{M}_M$. We may assume that f is perpendicular to $\mathfrak{M}_M$, or equivalently, that $b^*(t)f(t)/B^2(t+i0)$ is perpendicular to H^2. But $b^*(t)f(t)$ is in H^2, so $b^*(t)f(t)/B^2(t+i0)$ is also in H^2, and $f=0$. Hence $\mathfrak{M}_M=\{f\colon Sf=Mf\}$.

The proof that $\mathfrak{M}_m=\{f\colon Sf=mf\}$ when Ci) and ii) hold follows similarly.

4. Index of a subset of *R*. We shall define a function I on the Lebesgue measurable subsets of R to the non-negative integers or ∞ that we will use later on in order to discuss the multiplicity of points of the continuous spectrum of S. We map the one point compactification R^* of R onto the unit circle U by means of the mapping $M: t \rightarrow e^{i\phi} = (t - i)/(t + i)$, so $t = \infty$ corresponds to $e^{i\phi} = 1$. Clearly intervals on R go into arcs on U. Suppose Γ is a Lebesgue measurable subset of R and Γ_U is the image of Γ under M. Then Γ_U satisfies exactly one of the following alternatives almost everywhere:

$A(0)$, Γ_U is empty or $\Gamma_U = U$; or

$A(n)$, Γ_U is the union of n disjoint closed non-degenerate arcs, where n is a positive integer; or

$A(\infty)$, neither of the above hold.

We define the *index* $I(\Gamma)$ of Γ to be y in case $A(y)$ holds.

For example, in case Γ is a bounded set, $I(\Gamma) = n$ if and only if Γ is a.e. equal to the union of n disjoint non-degenerate closed intervals. If $\Gamma = (-\infty, -1) \cup (1, \infty)$, then $I(\Gamma) = 1$.

We analytically describe $I(\Gamma)$ in

THEOREM 4. *The correspondence between triples* $(\mu[\infty], \beta, \mu)$ *and pairs* $(\alpha, \{f\})$ *of* (3c) *is such that* μ *is a singular measure if and only if* f *is a.e. equal to the indicator function of a measurable set* Γ, *and then* $I(\Gamma) = \dim L^2(d\mu) + \operatorname{sgn} \mu[\infty]$, *where* $\dim L^2(d\mu)$ *is the dimension of* $L^2(d\mu)$.

Proof. [2], Theorem [X], p. 386, shows that μ is singular if and only if f is an indicator function.

We shall use the mapping M to transform the situation of Theorem 4 to a problem on the unit circle U, where the index $I(\Gamma_U)$ of Γ_U was studied in [13]. There $I(\Gamma_U)$ was defined so that $I(\Gamma_U) = I(\Gamma)$. M induces a correspondence between measures μ on R satisfying $\int (1 + t^2)^{-1} d\mu(t) < \infty$ and finite measures μ_U on U having no point mass at $e^{i\phi} = 1$. It is given by the differential expression $2(t^2 + 1)^{-1} d\mu(t) = d\mu_U(\phi)$. For $\operatorname{Im} \lambda > 0$ we put $z = (\lambda - i)/(\lambda + i)$ and $\Phi_U(z) = \Phi(\lambda)$. Then (3.2) and (3.1) become

$$\Phi_U(z) = -i \exp[\alpha + \tfrac{1}{2} i \int_{\Gamma_U} (e^{i\phi} + z)/(e^{i\phi} - z)\, d\phi] \tag{4.1}$$

$$\Phi_U(z) = i(1 + z)/(1 - z)\mu[\infty] + \tfrac{1}{2} i \int (e^{i\phi} + z)/(e^{i\phi} - z)\, d\mu_U(\phi), \quad |z| < 1. \tag{4.2}$$

It follows from (4.1) and (4.2) and Theorems 1 and 2 of [13] that $I(\Gamma) = I(\Gamma_U) = \dim L^2(d\mu_U)$, which equals $\dim L^2(d\mu) + \operatorname{sgn}\mu[\infty]$.

Some light is shed on the role of the $\operatorname{sgn}\mu[\infty]$ term in Theorem 4 by considering $\Gamma = (-\infty, 0)$. Then, if $\alpha = 0$, one finds $\Phi(\lambda) = \lambda$, $\mu[\infty] = 1$, $\mu = 0$ and $I(\Gamma) = 1$. The $\mu[\infty]$ term effectively adds a point mass at ∞ to the measure μ if $\mu[\infty] > 0$.

5. Diagonalization of S. We are now in a position to put the pieces together. We have the expression (2.4) for the resolvent of S and we know from Theorems 2 and 3 that S is AC on R except for the possibility of eigenvalues at m or M. Our main result, a diagonalization of S, is formulated in

THEOREM 5. *Let S be the self-adjoint operator on $L^2(E)$ specified in* (1.1), *and fix a measurable real-valued function α on R. For each real number ξ define the measurable set Γ_ξ and the function Φ_ξ by*

$$(5.1)\qquad \Gamma_\xi = \{t \in E : -\infty < (a(t) + |b(t)|^2 - \xi)/(a(t) - |b(t)|^2 - \xi) \leqq 0\}$$

and

$$\Phi_\xi(\lambda) = \exp(\alpha(\xi) + \int_{\Gamma_\xi} [(t-\lambda)^{-1} - t(t^2+1)^{-1}]dt), \quad \operatorname{Im}\lambda > 0.$$

Φ_ξ has a unique representation of the form

$$(5.2)\qquad \begin{aligned} \Phi_\xi(\lambda) = {} & \mu_\xi[\infty]\lambda + \beta(\xi) \\ & + \int [(t-\lambda)^{-1} - t(t^2+1)^{-1}] d\mu_\xi(t), \quad \operatorname{Im}\lambda > 0, \end{aligned}$$

where $\mu_\xi[\infty] \geqq 0$, $\beta(\xi)$ is real and μ_ξ is a measure such that $\int (t^2+1)^{-1} d\mu_\xi(t) < \infty$. Let $\mathcal{H}_\xi$ be the Hilbert space with inner product $\langle\cdot,\cdot\rangle_\xi$ defined by

$$\langle f, g\rangle_\xi \doteq \mu_\xi[\infty] f(\infty) g^*(\infty) + \int f(t) g^*(t) d\mu_\xi(t).$$

By this we mean that $\mathcal{H}_\xi$ is $L^2(d\mu_\xi)$ if $\mu_\xi[\infty] = 0$, otherwise $\mathcal{H}_\xi$ is the direct sum of $L^2(d\mu_\xi)$ and a one dimensional Hilbert space. Let P_m and P_M be the orthogonal projections on the linear manifolds $\mathfrak{M}_m$ and $\mathfrak{M}_M$ respectively.

Then there exists a continuous direct sum $\int \mathcal{H}_\xi d\xi$ of the Hilbert spaces $\mathcal{H}_\xi$, $\xi \in R$ and a partial isometry U with initial set $(\mathfrak{M}_m \oplus \mathfrak{M}_M)^\perp$ in $L^2(E)$ and final set $\int \mathcal{H}_\xi d\xi$ such that

$$(5.3)\qquad \begin{aligned} \langle h(S)f, g\rangle = {} & \int h(\xi) \langle Uf(\xi), Ug(\xi)\rangle_\xi d\xi \\ & + h(m)\langle P_m f, g\rangle + h(M)\langle P_M f, g\rangle \end{aligned}$$

for all continuous bounded functions h on R.

U is defined on the total set $\{k_u\colon u$ *non-real,* $k_u(t) = b(t)/(t-u)\}$ *of* $L^2(E)$ *by*

$$(5.4)\qquad Uk_u(\xi) = [2^{\frac{1}{2}} \exp(\alpha(\xi)/2) W^*(u^*, \xi + i0)(t-u)]^{-1}$$

if $-\infty < t < \infty$ *and*

$$= [2^{\frac{1}{2}} \exp(\alpha(\xi)/2) W^*(u^*, \xi + i0)]^{-1} \text{ if } t = \infty.$$

For any $F \in \int \mathfrak{H}_\xi \, d\xi$,

$$(5.5)\quad b^*(x)(U^*F)(x) = \lim_{y\downarrow 0} \tfrac{1}{2\pi i} \int \langle F(\xi), Uk_{z^*}(\xi) - Uk_z(\xi) \rangle_\xi \, d\xi \text{ a. e.}$$

where $z = x + iy$, *with* x, y *real.*

The multiplicity $n(\xi)$ *of the point* ξ *of the continuous spectrum of* S *is equal to* $I(\Gamma_\xi)$.

Proof. We employ the Stieltjes inversion formula to find the almost everywhere existing derivative $E'(\xi)$ of $E(\xi) = \langle E_S(\xi) k_u, k_v \rangle$, where $E_S(\cdot)$ is now the resolution of the identity of S. Let $\lambda = \xi + i\eta$, with ξ and η real, $\eta > 0$. Now

$$E'(\xi) = \tfrac{1}{2\pi i} \lim_{\eta\downarrow 0} [\langle (S-\lambda)^{-1} k_u, k_v \rangle - \langle (S-\lambda)^{-1} k_v, k_u \rangle^*],$$

which by Theorem 1

$$= \lim_{\eta\downarrow 0} \tfrac{1}{2} [(v^* - u) W^*(u^*, \lambda) W(v^*, \lambda)]^{-1} \\ [W(v^*, \lambda) W^*(v, \lambda) - W(u, \lambda) W^*(u^*, \lambda)].$$

By (3.5) and then (5.2) this equals

$$[2(v^* - u) \exp \alpha(\xi) W^*(u^*, \xi + i0) W(v^*, \xi + i0)]^{-1} [\Phi_\xi^*(v) - \Phi_\xi(u)]. \\ = \langle Uk_u(\xi), Uk_v(\xi) \rangle_\xi.$$

From this and Theorem 2 we deduce that (5.3) holds for all continuous bounded functions h provided $f = k_u$ and $g = k_v$. The existence of the direct integral now follows, since μ_ξ is a singular measure for each $\xi \in R$, by duplicating the proof of Theorem 3 of [13].

For almost all x in E

$$b^*(x)(U^*F)(x) = \lim_{y\downarrow 0} \langle U^*F, b \cdot P_z \rangle \\ = \lim_{y\downarrow 0} \tfrac{1}{2\pi i} \langle U^*F, k_{z^*} - k_z \rangle \\ = \lim_{y\downarrow 0} \tfrac{1}{2\pi i} \int \langle F(\xi), Uk_{z^*}(\xi) - Uk_z(\xi) \rangle_\xi \, d\xi,$$

which establishes (5.5).

If $n(\xi)$ is the multiplicity of the point ξ, then

$$n(\xi) = \dim L^2(d\mu_\xi) + \operatorname{sgn} \mu_\xi[\infty],$$

which by Theorem 4, $= I(\Gamma_\xi)$.

COROLLARY. i) *Suppose $R - E$ is not of measure zero. Then the spectrum of S is $[m, M]$.*

ii) *Suppose $R = E$. Then the spectrum of S is $[m, m_1] \cup [M_1, M]$, where $m_1 = \operatorname{ess\,sup}(a - |b|^2)$ and $M_1 = \operatorname{ess\,inf}(a + |b|^2)$.*

Proof. This follows from a consideration of $n(\xi) = I(\Gamma_\xi)$. Thus we see that the spectrum of S has either one or two components when $E = R$.

COROLLARY. *Suppose $(\mu[\infty], \beta, \mu)$ and $(\alpha, \{f\})$ are in correspondence as in Theorem 4, with μ a singular measure and f the indicator function of a measurable set Γ. Assume $R \backslash \Gamma$ has positive measure. Consider the operator S_0 on $L^2(\Gamma)$ defined by*

$$S_0 g(x) = (\pi i)^{-1} \int_\Gamma^{*} (t - x)^{-1} g(t)\, dt.$$

Define $\int_{-1}^{1} \mathfrak{H}_\xi\, d\xi$ to be the Lebesgue space with inner product

$$\langle f, g \rangle = \mu[\infty] \int_{-1}^{1} f(\xi, \infty) g^*(\xi, \infty)\, d\xi$$
$$+ \int_{-1}^{1} \int f(\xi, t) g^*(\xi, t)\, d\mu(t)\, d\xi.$$

Then there is a unitary mapping U_0 of $L^2(\Gamma)$ onto $\int_{-1}^{1} \mathfrak{H}_\xi\, d\xi$ such that U_0 diagonalizes S_0, that is, $US_0U^{-1}f(\xi, t) = \xi f(\xi, t)$ in $\int_{-1}^{1} \mathfrak{H}_\xi\, d\xi$. Uk_u is defined by (5.4) with $\alpha(\xi) = \alpha$, $a = 0$, and $b = 1$.

Proof. Set $a = 0$, $b = 1$ in Theorem 5.

6. Commutators with a one dimensional range. Suppose S, as defined by 1.1, is a bounded operator and E is a bounded set. Define the operator A by demanding that A send any $f(t)$ in $L^2(E)$ into $tf(t)$ in $L^2(E)$. Clearly A has simple spectra with cyclic vector b, that is,

$$\{A^n b : 0, 1, 2, \cdots\} \text{ is total in } L^2(E).$$

Furthermore $SA - AS = (\pi i)^{-1} \langle \cdot, b \rangle b$. Conversely, the following result can be easily deduced from the paper [4] of Xa Dao-Xeng:

THEOREM 6. *Suppose* $\mathcal{A}$ *and* $\mathcal{S}$ *are bounded self-adjoint operators on a complex Hilbert space* $\mathcal{H}$ *and* e *is in* $\mathcal{H}$. *Suppose*

$$\mathcal{S}\mathcal{A}-\mathcal{A}\mathcal{S}=(\pi i)^{-1}\langle\cdot, e\rangle e$$

and $\{\mathcal{A}^n e\colon n=0,1,2,\cdots\}$ *is total in* $\mathcal{H}$.

Then there exists a Lebesgue measurable bounded subset E *of* R *and a unitary map* T *of* $\mathcal{H}$ *onto* $L^2(E)$ *such that* $T\mathcal{A}T^{-1}=A$ *and* $T\mathcal{S}T^{-1}=S$, *where* $Af(t)=tf(t)$ *for all* $f\in L^2(E)$ *and* S *is of the form* (1.1).

Proof. We shall prove the above result by showing it to be a corollary of the following theorem of C. R. Putnam's (see [10]):

Let $\mathcal{A}$ *be a bounded self-adjoint operator and* $\mathcal{B}$ *a bounded operator in a Hilbert space* $\mathcal{H}$ *such that* $\mathcal{B}\mathcal{A}-\mathcal{A}\mathcal{B}=(\pi i)^{-1}C$, *where* C *is a non-negative operator. Let* $E_{\mathcal{A}}$ *be a resolution of the identity of* $\mathcal{A}$. *Then* $\langle E_{\mathcal{A}}(\cdot)Cx, Cx\rangle$ *is absolutely continuous for every* x *in* $\mathcal{H}$.

We shall use Putnam's result since we feel that it points the way towards a fruitful generalization of the theory in this paper.

Assume the hypotheses of Theorem 6. Then by Putnam's theorem $\langle E_{\mathcal{A}}(\cdot)e, e\rangle$ is absolutely continuous, so $d\langle E_{\mathcal{A}}(t)e, e\rangle/dt$ exists and is non-negative a.e.. Let $E=\{t\colon d\langle E_{\mathcal{A}}(t)e, e\rangle/dt>0\}$ and selsct $\omega(\cdot)$ to be a measurable function such that $|\omega(t)|=1$ for all t in E. Put

$$b(t)=\omega(t)\,[d\langle E_{\mathcal{A}}(t)e, e\rangle/dt]^{\frac{1}{2}}.$$

Define the linear operator T_0 by specifying that $(T_0\mathcal{A}^n e)(t)=t^n b(t)$ for $n=0,1,2,\cdots$. Then for any polynomial p,

$$\begin{aligned}\|T_0 p(\mathcal{A})e\|^2&=\int |p(t)|^2 d\langle E_{\mathcal{A}}(t)e, e\rangle\\&=\int_E |p(t)b(t)|^2\,dt,\end{aligned}$$

so T_0 has a unique isometric extension to a unitary operator T on $\mathcal{H}$ onto $L^2(E)$ and $T\mathcal{A}T^{-1}=A$. Let $S=T\mathcal{S}T^{-1}$, so

$$SA-AS=(\pi i)^{-1}\langle\cdot, b\rangle b. \tag{6.1}$$

We shall show that all bounded self-adjoint solutions S of (6.1) are of the form (1.1). Clearly any bounded S of the form (1.1) satisfies (6.1). If S_2 is any other solution, then $S-S_2$ commutes with A, and since A has simple spectra, is necessarily a function of A ([1], p. 220, Satz 1). Thus S is of form (1.1).

THEOREM 7. *Suppose $\mathcal{A}$ and $\mathcal{S}$ satisfy the hypotheses of Theorem* 6. *Let $z=x+iy$, with x and y real. Then for any complex number λ with* $\operatorname{Im}\lambda>0$,

$$w(x,\lambda)-1=\lim_{y\downarrow 0}\frac{2}{\pi}\,y\langle(\mathcal{S}-\lambda)^{-1}(\mathcal{A}-z)^{-1}e,(\mathcal{A}-z)^{-1}e\rangle \tag{6.2}$$

exists for almost all x. w is a.e. *of the form* (2.2) *where a is a real and b a complex $L^{\infty}(E)$ function.*

$\mathcal{S}$ is AC on R and the spectral multiplicity function n of S is given by $n(\xi)=I(\Gamma_{\xi})$, where Γ_{ξ} is defined by (5.1). *Specification of Lebesgue measure on $[m,M]$ and n gives a complete set of unitary invariants for $\mathcal{S}$.*

Proof. Due to the unitary mapping T it suffices to prove that

$$w(x,\lambda)-1=\lim_{y\downarrow 0}\frac{2}{\pi}\,y\langle(\mathcal{S}-\lambda)^{-1}k_{z},k_{z}\rangle$$

in order to derive (6.2). This, however, is a consequence of (2.4) and (2b). The remaining statements are deducible from Theorem 5.

UNIVERSITY OF VIRGINIA.

REFERENCES.

[1] N. I. Achieser and I. M. Glasmann, *Theorie der linearen operatorem im Hilbertraum*, Akademie Verlag, Berlin (1954).

[2] N. Aronszajn and W. F. Donoghue, Jr., "On exponential representations of analytic functions in the upper half-plane with positive imaginary part," *Journal D'Analyse Mathématique*, vol. 5 (1956-57), pp. 321-388.

[3] A. Calderón, F. Spitzer and H. Widom, "Inversion of Toeplitz matrices," *Illinois Journal of Mathematics*, vol. 3 (1959), pp. 490-498.

[4] Xa Dao-xeng, "On non-normal operators," *Chinese Mathematics* (American Mathematical Society translation of Acta Mathematica Sinica), vol. 3 (1963), pp. 232-246.

[5] W. Koppelman and J. Pincus, "Spectral representations for finite Hilbert transformations," *Mathematische Zeitschrift*, vol. 71 (1959), pp. 399-407.

[6] W. Koppelman, "On the spectral theory of singular integral operators," *Transactions of the American Mathematical Society*, vol. 97 (1960), pp. 35-63.

[7] . G. Krein, "Integral equations on the half line with kernels depending on the difference of the arguments," *American Mathematical Society Translations*, series 2, vol. 22, pp. 163-288.

[8] V. I. Krylov, "On functions regular in a half-plane," *American Mathematical Society Translations*, vol. 32 (1963), pp. 37-81.

[9] J. D. Pincus, "On the spectral theory of singular integral operators," Brookhaven National Laboratory AMD 313, BNL 6915. To appear.

[10] C. R. Putnam, "Commutators, absolutely continuous spectra, and singular integral operators," *American Journal of Mathematics*, vol. 86 (1965), pp. 310-316.

[11] M. Rosenblum, "The absolute continuity of Toeplitz's matrices," *Pacific Journal of Mathematics*, vol. 10 (1960), pp. 987-996.

[12] ———, "Self-adjoint Toeplitz operators and associated orthonormal functions," *Proceedings of the American Mathematical Society*, vol. 13 (1963), pp. 590-595.

[13] ———, "A concrete spectral theory for self-adjoint Toeplitz operators," *American Journal of Mathematics*, vol. 87 (1965), pp. 709-718.

[14] F. G. Tricomi, *Integral equations*, Interscience, New York, 1957.

[15] H. Widom, "Singular integral equations in L_p," *Transactions of the American Mathematical Society*, vol. 97 (1960), pp. 131-160.

[16] W. Koppelman, "Spectral multiplicity theory for a class of singular integral operators," *Transactions of the American Mathematical Society*, vol. 113 (1964), pp. 87-100.

[17] J. D. Pincus, "On the spectral theory of singular integral operators," *Transactions of the American Mathematical Society*, vol. 13 (1964), pp. 101-129. (Published version of [9]).

[18] ———, "Commutators, generalized eigenfunction expansions and singular integral operators," Brookhaven National Laboratory report BNL 8668.

[19] C. R. Putnam, "The spectra of generalized Hilbert transforms," *Journal of Mathematics and Mechanics*, vol. 14 (1965), pp. 857-872.

INVARIANT DOMAINS FOR KLEINIAN GROUPS.[1]

By R. D. M. Accola.

Introduction. Let Γ be a properly discontinuous group of fractional linear transformations of the Riemann sphere, S.[2] Let Ω be the set of discontinuity and let Σ be the limit set. Let $\{\Omega_i\}$ be an enumeration of the components of Ω. Let $\Gamma(\Omega_i)$ be the subgroup of those elements of Γ which map Ω_i onto itself. More generally, if O is a (non-empty) domain, i.e., a connected open set, let $\Gamma(O)$ be the subgroup of those elements of Γ which map O onto itself. If $\Gamma(O) = \Gamma$ (or $\Gamma(\Omega_i) = \Gamma$), call $O(\Omega_i)$ an *invariant domain* (*invariant component*). If $\Gamma(\Omega_i)$ is merely the identity transformation, call Ω_i an *atom.* If Γ is a group which leaves a simply connected domain O with at least two boundary points invariant, then one can map O onto the unit disk, and in a natural way Γ is carried into a Fuchsian group Γ'. We will describe this by saying that Γ acting on O is equivalent to a Fuchsian group. If Γ' is a Fuchsian group of the first (or second) kind, we will say that Γ acting on O is equivalent to a Fuchsian group of the first (or second) kind.

By a Kleinian group we mean a group where Σ has three or more distinct points. Otherwise a group will be called elementary. The object of this paper is to study Kleinian groups with invariant domains. The best known examples of such groups are the Fuchsian groups where the inside and outside of some circle are invariant domains. The main results of this paper can be summarized by saying that, topologically, the Fuchsian groups are almost typical of Kleinian groups with two or more invraiant domains. In particular, it is proved that only cyclic groups have more than two invariant domains. If a Kleinian group has two invariant domains, one can find a pair of maximal invariant domains, each of which is simply connected. In this case,

---

Received August 13, 1964.

[1] The research for this paper was sponsored by the Office of Naval Research, Contract number: Nonr 562(31). The author wishes to thank Professors L. Ahlfors and L. Greenberg for their many helpful remarks on the subject of this paper.

[2] For a discussion of this subject see L. R. Ford, *Automorphic Functions.* We shall use the term loxodromic to describe non-elliptic transformations with two fixed points. The term then will include what are usually called hyperbolic transformations.

if there are non-invariant components of the set of discontinuity, they are atoms. The final result of the paper is really a corollary of a deep theorem of Nielsen and Fenchel. If a group Γ has a maximal pair, O_1 and O_2, of invariant domains, then the quotient surfaces O_1/Γ and O_2/Γ are homeomorphic. In the last section we give some examples.

Main results. The proofs of the main theorems will be broken into several lemmas.

LEMMA 1. *Suppose O is an invariant domain for a cyclic loxodromic group generated by T. Then there exists a Jordan arc γ connecting the fixed points of T which lies entirely within O with the possible exception of one or both of the fixed points.*

Proof. If p is in O, let σ be a path lying in O connecting p to Tp. Then $\gamma' = \bigcup_{-\infty}^{\infty} T^n\sigma$ is a continuous image of the open unit interval lying entirely within O. Since $\lim_{n\to+\infty} T^n\sigma$ is one fixed point of T and $\lim_{n\to-\infty} T^n\sigma$ is the other fixed point of T, we may add these fixed points to γ' to obtain γ'' a continuous image of the closed unit interval. We can then find in γ'' a homeomorphic image, γ, of the closed unit interval with the desired properties. q. e. d.

LEMMA 2. *Suppose O is a domain invariant under a cyclic parabolic group generated by T. Then there exists a closed Jordan curve γ, one of whose points is the fixed point of T, lying entirely within O with the possible exception of the fixed point.*

Proof. Exactly the same as the proof of Lemma 1.

Notice that in both Lemma 1 and Lemma 2 the curve γ is a subset of a set invariant under the cyclic group. If in Lemmas 1 and 2 O is assumed to be simply connected with at least two boundary points, one may assume that γ is actually invariant under T, for in this case, the cyclic group acting on O is equivalent to a cyclic Fuchsian group. In all applications of these lemmas, O will be simply connected.

LEMMA 3. *Suppose O is a simply connected domain with at least two boundary points, invariant under a finite cyclic elliptic group generated by T. Then exactly one of the fixed points of T lies in O.*

Proof. O is conformally equivalent to the unit disk, D, and under this conformal map T is carried into the generator T' of a finite cyclic Fuchsian group. T' is, therefore, elliptic and has one fixed point in D. The result now follows. q.e.d.

The following familiar result is included for future reference.

LEMMA 4. *If A is a closed connected set, then every component of $S-A$ is simply connected.*

Definitions. For a properly discontinuous group, Γ, an ordered pair (triple) of mutually disjoint invariant (non-empty) domains will be called an *invariant pair* (*invariant triple*). An invariant pair (O_1, O_2) will be called *maximal* if, whenever we have another invariant pair (O_1', O_2') such that $O_i \subset O_i'$, then $O_i = O_i'$, $i=1,2$. There is a similar definition of a maximal invariant triple.

Note that for an invariant pair, $\Sigma \subset \partial O_1 \cap \partial O_2$.

LEMMA 5. *Let A be a closed invariant set for Γ. Let O be an invariant domain disjoint from A. Let O' be the component of $S-A$ containing O. Then O' is invariant.*

Proof. If T is in Γ, T maps components of $S-A$ onto components of $S-A$. Since $TO' \cap O'$ is non-empty, TO' must equal O'. q.e.d.

LEMMA 6. *Let (O_1, O_2) be an invariant pair for Γ. Then there exist maximal invariant pairs. If (O_1, O_2) is a maximal invariant pair, then O_1 and O_2 are simply connected.*

Proof. The existence of maximal invariant pairs is a standard application of Zorn's lemma. Also several applications of Lemma 5 can be used for an explicit construction. To show O_2 simply connected we note that O_2 must equal the component of $S-\bar{O}_1$ containing it.

LEMMA 7. *Let (O_1, O_2) be a maximal invariant pair. Let T be a parabolic transformation in Γ such that $Tz = z+1$. Then there exists a real number M such that the domains $\{y > M\}$ and $\{y < -M\}$ lie in O_1 and O_2, or O_2 and O_1, respectively.*

Proof. By Lemma 2 and the simple connectivity of O_1 and O_2 there are curves γ_1 and γ_2 invariant under T lying in O_1 and O_2 respectively, each of which divides the plane. Thus the two curves divide the plane into three regions. Let U be the region bounded by γ_1, γ_2, and the point at infinity. Then $\Sigma \subset U$. For if a non-elliptic fixed point lies in the Jordan region determined by, say, γ_1, then a point of O_2 lies in this Jordan region. Thus this point can be joined to a point on γ_2 by a path in O_2 not intersecting γ_1, a contradiction. By the maximality of O_1 it follows that the entire Jordan region determined by γ_1 lies in O_1.

Since each γ_i is a path invariant under T, there is some number M such that each path lies in the strip $\{-M < y < M\}$. The result now follows.

THEOREM 1. *Let* Γ *be a properly discontinuous group with an invariant triple. Then* Γ *is cyclic.*

Proof. As in Lemma 6, there exist maximal invariant triples. Let (O_1, O_2, O_3) be a maximal invariant triple. We show first that at least two of the domains in this triple are simply connected. Suppose O_1 is not simply connected. If O_2 and O_3 lie in the same component of $S - \bar{O}_1$, then O_1 is not maximal. Thus O_2 and O_3 lie in different components of $S - \bar{O}_1$. By Lemma 5, O_2 and O_3 are components of $S - \bar{O}_1$ and by Lemma 4, they are simply connected.

To show that Γ is cyclic, we examine the possibilities for Σ. If Σ is empty, Γ is a finite group and is, therefore, cyclic since it is isomorphic to a finite Fuchsian group.

Assume now that Σ is non-empty. Then each O_i is simply connected, for O_1 is a component of $S - \bar{O}_2 \cup \bar{O}_3$ and $\bar{O}_2 \cup \bar{O}_3$ is connected since $\Sigma \subset \partial O_2 \cap \partial O_3$. Since an elliptic element of Γ would have to have a fixed point in each O_i, there are no elliptic elements in Γ. To show that Γ is an elementary group we show that if T is a loxodromic element in Γ, then any other element of Γ has its fixed points among those of T. By Lemma 1 there are Jordan arcs γ_1, γ_2, γ_3 connecting the fixed points of T lying in O_1, O_2, O_3 respectively, except for the endpoints. By the Jordan curve theorem these arcs divide S into three domains U_{12}, U_{13} and U_{23}, where $\partial U_{ij} = \gamma_i \cup \gamma_j$. Suppose that T' is an element of Γ with a fixed point p not among those of T. Since p lies on no arc γ_i, p lies in one of the three regions U_{ij}, say U_{12}. Since $p \in \partial O_3$, there is a point q of O_3 lying in U_{12}. γ_3 is exterior to U_{12}. Since we may join q to γ_3 by a path lying entirely in O_3, we have reached the desired contradiction.

Thus if Σ is non-empty, Γ is a cyclic loxodromic group or a purely parabolic group. In the latter case it is abelian and, therefore, a cyclic group since, again, it is isomorphic to a Fuchsian group. All possible cases have now been examined.[3] q. e. d.

---

[3] L. R. Ford, *ibid.* Chapter VI.

THEOREM 2. *Suppose* Γ *is a Kleinian group with an invariant domain* O_1. *Let* O_2 *and* Ω_3 *be domains such that the three domains are mutually disjoint and* Ω_3 *is a component of the set of discontinuity. Then* $\Gamma(O_2) \cap \Gamma(\Omega_3)$ *is non-loxodromic. If, in addition,* O_1 *is simply connected, then* $\Gamma(O_2) \cap \Gamma(\Omega_3)$ *is non-elliptic.*

Proof. Since O_1 is invariant for Γ, the component O_1' of $S - \Sigma$ containing O_1 is also invariant and so Ω_3, a component of $S - \bar{O}_1'$, is simply connected. Also, since Γ is Kleinian, $\bar{O}_1 \cup \Omega_3$ is connected. If O_2' is the component of $S - \bar{O}_1 \cup \bar{\Omega}_3$ containing O_2, it is simply connected and invariant under $\Gamma' = \Gamma(O_2) \cap \Gamma(\Omega_3)$. If, in addition, O_1 is simply connected, any elliptic element in Γ' must have a fixed point in O_1, O_1' and Ω_3. This is a contradiction.

Suppose T is a loxodromic element in Γ'. By Lemma 1 there are Jordan arcs γ_2 and γ_3 lying in O_2' and Ω_3 respectively which divide the plane into two regions, each of which contain boundary points of Ω_3. Thus each of these regions contains points of O_1, which is invariant. Since $O_1 \cap (\gamma_2 \cup \gamma_3) = \emptyset$, we have a contradiction to the connectivity of O_1. q. e. d.

As an application of Theorem 2, we observe that if a Kleinian group has one invariant component Ω_0, then $\Gamma(\Omega_i) \cap \Gamma(\Omega_j)$, $0 \neq i \neq j \neq 0$, is either the identity, a cyclic parabolic group or a cyclic elliptic group. Easy examples show that all these cases can occur.

THEOREM 3. *Suppose* Γ *is a Kleinian group with an invariant pair. Then every non-invariant component of the set of discontinuity is an atom.*

Proof. Let (O_1, O_2) be a maximal invariant pair. Let Ω_3 be a component of Ω disjoint from O_1 and O_2. Then $\Gamma(\Omega_3) = \Gamma(O_1) \cap \Gamma(O_2) \cap \Gamma(\Omega_3)$ and O_1 is simply connected, so by Theorem 2 we need only show that $\Gamma(\Omega_3)$ cannot contain a parabolic element. Suppose T, $Tz = z + 1$, leaves Ω_3 fixed. We argue as before. There are three Jordan curves $\gamma_1, \gamma_2, \gamma_3$, invariant under T lying in O_1, O_2, Ω_3 respectively, dividing the plane. We may assume, by Lemma 7, that γ_1 and γ_2 are the lines $\{y = M\}$ and $\{y = -M\}$. Since γ_3 is invariant under T, γ_3 divides the strip $\{-M < y < M\}$ into 2 regions. Since $\partial\Omega_3 \subset \Sigma$, there is a limit point of the group Γ lying in the region between γ_1 and γ_3. Thus there is a point of O_2 in this same region and we arrive at a contradiction as before. q. e. d.

A group with an invariant pair may be considered a generalization of a Fuchsian group. For such groups there are either two or one invariant components, cases analogous to Fuchsian groups of the first or second kind. We give an example below to show that atoms can occur in such groups.

We are grateful to Leon Greenberg for pointing out how the following theorem is a consequence of a theorem of Nielsen and Fenchel.

THEOREM 4. *Let* (O_1, O_2) *be a maximal invariant pair for a Kleinian group* Γ. *Then the surfaces* O_1/Γ *and* O_2/Γ *are homeomorphic.*

Proof. Let ϕ_i be a conformal map of O_i onto the unit disk D, $i = 1, 2$. Let ψ_i map Γ onto the corresponding Fuchsian group Γ_i. There is a natural isomorphism, τ, of Γ_1 onto Γ_2 given by $\tau\psi_1 T = \psi_2 T$ for $T \in \Gamma$. Suppose that T is loxodromic. Lemma 1 insures that the two fixed points of T are distinct accessible boundary points for both O_1 and O_2. By boundary correspondence theorems concerning accessible boundary points, $\psi_1 T$ and $\psi_2 T$ are hyperbolic. If T is parabolic, Lemma 7 insures that the fixed point of T defines only one accessible boundary point of O_1 and O_2, and so $\psi_1 T$ and $\psi_2 T$ are parabolic. Thus τ maps parabolic elements into parabolic elements and hyperbolic elements into hyperbolic elements. Moreover, by the methods used above, it follows that if $\psi_1 T$ and $\psi_1 S$ are hyperbolic elements of Γ_1, whose axes cross, then the same is true for the axes of $\psi_2 T$ and $\psi_2 S$. Thus the isomorphism τ is a geometric isomorphism in the sense of Nielsen and Fenchel, and so D/Γ_1 is homeomorphic to D/Γ_2.[4]

Examples.[5] The following easy lemma gives a useful criterion for determining when a component of a fundamental region lies in an invariant component of the set of discontinuity.

LEMMA 8. *Let* Γ *be a Kleinian group. Let* F *be a component of a fundamental region for* Γ *and let* p *be a point of* F. *Suppose for each transformation,* T_i, *in a set of generators* $\{T_i\}$ *of* Γ, *it is true that* $T_i p$ *can be joined to* p *by a path lying in* Ω. *Then the component of* Ω *containing* F *is invariant.*

Example 1. Ω and Σ are both connected. A sequence of positive numbers

---

[4] J. Nielsen and W. Fenchel, *Discontinuous Groups of Non-Euclidean Motions.* Chapter V. A precise statement of the theorem of Nielsen and Fenchel is as follows. Let F and F' be isomorphic Fuchsian groups. Suppose the isomorphism carries parabolic transformations into parabolic transformations and hyperbolic transformations into hyperbolic transformations. Moreover, suppose that the axes of two hyperbolic elements in F intersect if and only if the axes of the corresponding elements in F' intersect. Then the quotient surfaces of the unit disk modulo F and F' are homeomorphic surfaces. (The groups need not be finitely generated.)

[5] The following examples are elaborations of the ideas in L. R. Ford, *ibid.*, pp. 55-59.

a_n can be chosen inductively so that the circles in the double sequence of circles $\{C_{n,k}\colon n=1,2,\cdots;k=1,3,5,\cdots,2^n-1\}$

$$C_{n,k}\colon |z-k2^{-n}+ia_n|=a_n$$

are mutually disjoint. Necessarily $a_n\to 0$. If $C_{n,k}'$ is the reflection of $C_{n,k}$ in the real axis, let $T_{n,k}$ be the parabolic transformation whose isometric circles are $C_{n,k}$ and $C_{n,k}'$. If F is the exterior of all the circles $C_{n,k}$ and $C_{n,k}'$, then F is the fundamental region for the group generated by the $T_{n,k}$. Since F is the fundamental region, Ω is connected. But Ω is simply connected, since F is simply connected and $\bar{F}\cap\overline{T_{n,k}F}-\Sigma$ is homeomorphic to the real line for any $T_{n,k}$. Since Ω is simply connected, it follows that Σ is connected.

Example 2. Γ has two invariant components and atoms. Let S_1 be an infinite spiral in the annulus $\{1<z<2\}$ which converges to both the circles $\{|z|=1\}$ and $\{|z|=2\}$. Let R be the annulus minus the spiral, a simply connected region. Let S_2 be another infinite spiral with the same properties as S_1 which divides R into two simply connected domains, R_1 and R_2. Notice that $\partial R_i=S_1\cup S_2\cup\{|z|=1\}\cup\{|z|=2\}$, $i=1,2$. Now replace each spiral, S_i, by a sequence of circles $\{C_{i,k}\colon k=0,\pm 1,\pm 2,\cdots\}$, $i=1,2$, so that a) $C_{i,k-1}$ is tangent to $C_{i,k}$ at a point on S_i, $i=1,2$; and b) $C_{1,k}\cap C_{2,j}=\emptyset$ for all k and j; c) the exterior of all the circles is divided into four domains $\{|z|<1\}$, $\{|z|>2\}$, F_1 and F_2, where F_i is what remains of R_i after the circles and their interiors have been removed. Let Γ' be the group generated by reflections in all the circles $C_{i,k}$, $i=1,2$, $k=0,\pm 1,\pm 2,\cdots$. Let Γ be the subgroup of index two of fractional linear transformations in Γ'. For Γ' F_1 and F_2 satisfy the criterion of Lemma 8 and lie in different components of Ω.[6] Also the regions $\{|z|<1\}$ and $\{|z|>2\}$ are atoms. Therefore, Γ has two invariant components and atoms.

The same methods can be used to produce a Kleinian group, Γ, with two invariant components, Ω_1 and Ω_2, having the following properties: when Ω_1 and Ω_2 are mapped onto the unit disk and Γ is carried into Fuchsian groups

---

[6] To show that F_1 and F_2 lie in different components of Ω we argue as follows. Let C be a generic notation for a $C_{i,k}$ or one of its translates under any T in Γ'. Suppose p_1 and p_2, lying in F_1 and F_2 respectively, can be joined by an arc γ (which can be assumed to be analytic) lying in Ω. Then γ intersects only a finite number of the C's. (Otherwise a limit point would lie on C.) Thus γ intersects the C's in a finite number of points. Suppose the number is N. We now proceed by induction on N. If $N=1$, then a connectivity argument, using essentially that F_1 is connected, shows that p_2 lies in a translate of F_1, which is a contradiction. For N greater than one, the case is easily reduced to the case $N=1$. Thus the existence of the arc γ leads to a contradiction and so the result follows.

Γ_1 and Γ_2, then Γ_1 is a group of the first kind and Γ_2 is a group of the second kind.

REFERENCES.

L. R. Ford, *Automorphic Functions*, Chelsea Publishing Company, New York, 2nd Edition, 1951.

J. Nielsen and W. Fenchel, *Discontinuous Groups of Non-Euclidean Motions*, Manuscript. (Not as yet published.)

A PRODUCT FORMULA FOR WALL'S OBSTRUCTION.

By Stephen M. Gersten.

0. Introduction. Let X be a CW complex of finite type; that is, X has only finitely many cells in each dimension. If in addition there is an integer $N > 2$ such that

i) $H^{N+i}(X;\mathcal{B}) = 0$ for all systems of local coefficients $\mathcal{B}$ and all $i > 0$, and

ii) $H_{N+i}(\tilde{X};\mathbf{Z}) = 0$, all $i > 0$, where $\tilde{X}$ is the universal cover of X, then Wall [4] has defined an obstruction* $\kappa_N(X)$ to finding a finite complex with the same homotopy type as X. He proves that $H_N(\tilde{X},\tilde{X}^{N-1};\mathbf{Z})$ is a finitely generated projective $\mathbf{Z}\pi_1(X)$ module and the obstruction $\kappa_N(X)$ is the class of this module in $K_0(\mathbf{Z}\pi_1(X))$.

We define a condition G_N on chain complexes $\mathcal{C} = \{C_n, d_n; n \geqq 0\}$ of finitely generated free left R modules C_n over a ring R, which enables us to define an element $\kappa_N(\mathcal{C}) \in K_0(R)$. If X is the complex above satisfying i) and ii) and $\mathcal{C} = \{C_n(\tilde{X})\}$, $C_n(\tilde{X})$ (integer chains) considered as a $\mathbf{Z}\pi_1(X)$ module, then $\mathcal{C}$ satisfies G_N and $\kappa_N(\mathcal{C}) = \kappa_N(X)$, Wall's obstruction.

The reason for studying complexes satisfying G_N is to set up the algebraic framework needed to prove the following.

Theorem 0.1. *If X is a CW complex of finite type which is dominated by a finite CW complex, and Y is a finite CW complex of dimension r, then $\kappa_N(X)$ and $\kappa_{N+r}(X \times Y)$ are defined and related by*

$$\kappa_{N+r}(X \times Y) = (-1)^r \chi(Y) \cdot i_*(\kappa_N(X)).$$

Here $i_*: K_0(\mathbf{Z}\pi_1(X) \to K_0(\mathbf{Z}\pi_1(X \times Y))$ is induced by $i: X \to X \times Y$ given by $i(x) = (x, y_0)$ with y_0 the base point of Y, and $\chi(Y)$ is the Euler-Poincaré characteristic of Y.

As an immediate consequence, we have

Corollary. *If X is a CW complex of finite type which is dominated*

Received March 4, 1965.

* Wall calls his obstruction $\theta(X)$, but we prefer to emphasize the dependence on N.

by a finite complex, then $X \times S^{2n+1}$ *has the homotopy type of a finite CW complex.*

This corollary for the case $n = 0$ was proved by M. R. Mather before Wall's obstruction was discovered.

1. Let R be a ring with unit and $\mathcal{C} = \{C_n, d_n; n \geqq 0\}$ be a chain complex of finite type over R. Thus, each C_i is a finitely generated free left R module. Let N be a non negative integer. We say $\mathcal{C}$ satisfies condition G_N if

1) $H_{N+i}(\mathcal{C}) = 0$, all $i > 0$, and
2) C_N/B_N is a projective R module, where B_N is the image of

$$d_{N+1} \colon C_{N+1} \to C_N.$$

LEMMA 1.1. *If* $\mathcal{C}$ *satisfies* G_N, *it also satisfies* G_{N+i}, $i \geqq 0$.

Proof. It suffices to show $\mathcal{C}$ satisfies G_{N+1}. Clearly $H_{N+1+i}(\mathcal{C}) = 0$ for all $i > 0$, so we need only verify that C_{N+1}/B_{N+1} is projective. From the exact sequence

$$0 \to B_N \to C_N \to C_N/B_N \to 0$$

we see that B_N is projective. From the exact sequence

$$0 \to Z_{N+1} \to C_{N+1} \to B_N \to 0$$

we see that $Z_{N+1} = \operatorname{Ker}(d_{N+1} \colon C_{N+1} \to C_N)$ is projective. But $H_{N+1} = Z_{N+1}/B_{N+1} = 0$, so $B_{N+1} = Z_{N+1}$. Thus,

$$C_{N+1}/B_{N+1} \approx B_N$$

is projective.

Definition 1.2. If $\mathcal{C}$ satisfies G_N, let $\kappa_N(\mathcal{C})$ be the class of C_N/B_N in $K_0(R)$, the group of stable isomorphism classes of finitely generated projective left R modules (1).

LEMMA 1.3. *If* $\mathcal{C}$ *satisfies* G_N, *then*

$$\kappa_{N+r}(\mathcal{C}) = (-1)^r \kappa_N(\mathcal{C}).$$

Proof. By induction it suffices to consider $r = 1$. In the proof of 1.1, we saw that

$$C_{N+1}/B_{N+1} \approx B_N.$$

Thus, there is an exact sequence

$$0 \to C_{N+1}/B_{N+1} \to C_N \to C_N/B_N \to 0,$$

whence $0 = \kappa_{N+1}(\mathcal{C}) + \kappa_N(\mathcal{C})$.

THEOREM 1.4. *If $\mathcal{C}$ satisfies G_N and $\kappa_N(\mathcal{C}) = 0$, thsn $\mathcal{C}$ is chain equivalent to a finite complex of free R modules.*

Proof. We may assume $N \geqq 1$ by 1.1. Let $D_N = C_N/B_N$. Since

$$0 \to B_N \to C_N \to D_N \to 0$$

splits, we may assume $C_N = B_N \oplus D_N$. Since $\kappa_N(\mathcal{C}) = 0$, there is a free R module R^n such that $D_N \oplus R^n$ is free. Let $\mathcal{C}'$ be the chain complex with

$$\begin{array}{ll} C'_N = D_N \oplus R^n & d'_N \mid D_N = d_N \mid D_N,\ d'_n \mid R^n = 1 \\ C'_{N-1} = C_{N-1} \oplus R^n & d'_{N-1} \mid C_{N-1} = d_{N-1},\ d'_{N-1} \mid R^n = 0 \\ C'_i = 0, i > N & d'_i = d_i, i < N-1 \\ C'_i = C_i, i < N-1 & \end{array}$$

and $\mathcal{C}''$ the complex with

$$\begin{array}{ll} C''_N = D_N; & d''_N = d_N \mid D_N \\ C'' = C_i, i < N; & d''_i = d_i, i < N. \\ C''_i = 0, i > N & \end{array}$$

There are natural inclusions which are chain maps

$$\begin{tikzcd} \mathcal{C}' & \mathcal{C}'' \arrow[l, "i"'] \arrow[r, "j"] & \mathcal{C} \end{tikzcd}$$

$\mathcal{C}' \xleftarrow{i} \mathcal{C}''$ is easily seen to be a chain equivalence (in fact it is an elementary expansion in the sense of Whitehead). One may easily show $\mathcal{C}'' \xrightarrow{j} \mathcal{C}$ is a chain equivalence using the fact $H_{N+i}(\mathcal{C}) = 0$, $i > 0$. Thus, $\mathcal{C}$ is chain equivalent to $\mathcal{C}'$, a finite free complex over R. This completes the proof.

We shall show later in 2.8 that the converse of 1.4 is valid.

It is convenient to give an alternative characterization of $\kappa_N(\mathcal{C})$. Let $\mathcal{C}^m$ denote the m-skeleton $(m \geqq 0)$ of $\mathcal{C}$. That is, $\mathcal{C}^m = \{C_n, d_n, 0 \leqq n \leqq m\}$. $\mathcal{C}^m$ is a subcomplex of $\mathcal{C}$ and the quotient complex

$$\mathcal{C}/\mathcal{C}^m = \{C_{n-1}, d_n; n \geqq m+2, d_{m+1} = 0\}.$$

PROPOSITION 1.5. *If $\mathcal{C}$ satisfies G_N, then $\kappa_N(\mathcal{C})$ is the class of $H_N(\mathcal{C}/\mathcal{C}^{N-1})$ in $K_0(R)$.*

Proof. $Z_N(\mathcal{C}/\mathcal{C}^{N-1}) = C_N$, $B_N(\mathcal{C}/\mathcal{C}^{N-1}) = B_N(\mathcal{C}) = B_N$. Thus

$$H_N(\mathcal{C}/\mathcal{C}^{N-1}) = C_N/B_N.$$

It is appropriate to point out the connection between $\kappa_N(\mathcal{C})$ and Wall's obstruction. Let X be a CW complex of finite type satisfying conditions i) and ii) of § 0. $(N > 2)$. From standard obstruction theory one sees that there is a map of pairs

$$f\colon (X, X^{N-1}) \rightarrow (X^N, X^{N-1})$$

with $f \mid X^{N-1} =$ identity, such that the composition

$$(X, X^{N-1}) \xrightarrow{f} (X^N, X^{N-1}) \xrightarrow{i} (X, X^{N-1})$$

is homotopic, through a homotopy of pairs, to the identity of (X, X^{N-1}). Here i is the inclusion. This implies that $\pi_N(X, X^{N-1})$ is a direct summand of $\pi_N(X^N, X^{N-1})$ as $\pi_1(X)$ modules. From the Hurewicz theorem $H_N(\tilde{X}, \tilde{X}^{N-1})$ is a direct summand of $H_N(\tilde{X}^N, \tilde{X}^{N-1})$ as $\pi_1(X)$ modules and since the latter is free, $H_N(\tilde{X}, \tilde{X}^{N-1})$ is projective. This is Wall's obstruction $\kappa_N(X)$. The chain complex $C_*(\tilde{X})$ satisfies G_N (as one sees from the proof of (1.5) and $\kappa_N(C_*(\tilde{X})) = \kappa_N(X)$, Wall's obstruction.

2. In this section we prove several algebraic lemmas that establish among other things the invariance of $\kappa_N(\mathcal{C})$ and G_N under chain equivalence.

Lemma 2.1. *Suppose we have a commutative diagram of R modules*

$$\begin{tikzcd}
0 \arrow[r] & C'_{N+2} \arrow[r, "i"] \arrow[d, "d'_{N+2}"] & C_{N+2} \arrow[r, "j"] \arrow[d, "d_{N+2}"] & C''_{N+2} \arrow[r] \arrow[d, "d''_{N+2}"] & 0 \\
0 \arrow[r] & C'_{N+1} \arrow[r, "i"] \arrow[d, "d'_{N+1}"] & C_{N+1} \arrow[r, "j"] \arrow[d, "d_{N+1}"] & C''_{N+1} \arrow[r] \arrow[d, "d''_{N+1}"] & 0 \\
0 \arrow[r] & C'_N \arrow[r, "i"] & C_N \arrow[r, "j"] & C''_N \arrow[r] & 0
\end{tikzcd}$$

where the rows are short exact sequences, the right column is exact at C''_{N+1} and $d_{N+1}d_{N+2} = 0$. Then the image sequence

$$0 \rightarrow B'_N \xrightarrow{i} B_N \xrightarrow{j} B''_N \rightarrow 0$$

is exact.

LEMMA 2.2. *Suppose we have a commutative diagram of R modules*

$$\begin{tikzcd}
& 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & B'_N \arrow[r] \arrow[d] & B_N \arrow[r] \arrow[d] & B''_N \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & C'_N \arrow[r] \arrow[d] & C_N \arrow[r] \arrow[d] & C''_N \arrow[r] \arrow[d] & 0 \\
& D'_N \arrow[r] \arrow[d] & D_N \arrow[r] \arrow[d] & D''_N \arrow[d] & \\
& 0 & 0 & 0 &
\end{tikzcd}$$

where the first two rows are exact and the columns are exact. Then the third row is a short exact sequence.

Lemmas 2.1 and 2.2 are proved by a rapid chase of diagrams.

THEOREM 2.3. *Suppose we have a short exact sequence of free R-chain complexes of finite type*

$$0 \to \mathcal{C}' \xrightarrow{i} \mathcal{C} \xrightarrow{j} \mathcal{C}'' \to 0.$$

Suppose further that $\mathcal{C}'$ satisfies $G_{N'}$ and $\mathcal{C}''$ satisfies $G_{N''}$. Then $\mathcal{C}$ satisfies G_N, $N = \max(N', N'')$ and

$$\kappa_N(\mathcal{C}) = \kappa_N(\mathcal{C}') + \kappa_N(\mathcal{C}'').$$

Proof. From the exact homology sequence one deduces that $H_{N+i}(\mathcal{C}) = 0$, $i > 0$. Let $D_N = C_N/B_N$, $D'_N = C'_N/B'_N$ and $D''_N = C''_N/B''_N$. Lemma 2.1 implies that the sequence

$$0 \to B'_N \to B_N \to B''_N \to 0$$

is exact. Lemma 2.2 then applies and proves that

$$0 \to D'_N \to D_N \to D''_N \to 0$$

is exact. Since D'_N and D''_N are projective, the sequence splits and $D_N \approx D'_N \oplus D''_N$. Thus C_N/B_N is projective, $\mathcal{C}$ satisfies G_N and

$$\kappa_N(\mathcal{C}) = \kappa_N(\mathcal{C}') + \kappa_N(\mathcal{C}'').$$

This completes the proof of the theorem.

Definition 2.4. If $\mathcal{C}$ is a chain complex, we define $\Sigma'\mathcal{C}$ to be the chain complex with $(\Sigma'\mathcal{C})_{n+1} = C_n$ and differential

$$(\Sigma'\mathcal{C})_{n+1} \xrightarrow{d_{n+1}} (\Sigma'\mathcal{C})_n$$

equal to the map $d_n\colon C_n \to C_{n-1}$. $\Sigma'\mathcal{C}$ is called the (first) suspension of $\mathcal{C}$. One defines inductively $\Sigma^{r+1}\mathcal{C} = \Sigma'(\Sigma^r\mathcal{C})$. We observe that if $\mathcal{C}$ satisfies G_N then $\Sigma^r\mathcal{C}$ satisfies G_{N+r} and

$$\kappa_{N+r}(\Sigma^r\mathcal{C}) = \kappa_N(\mathcal{C}).$$

Definition 2.5. Let $f\colon \mathcal{C}'' \to \mathcal{C}'$ be a chain map. The (reduced) mapping cylinder M_f of f is the chain complex with

$$(M_f)_n = C''_{n-1} \oplus C'_n \text{ and}$$
$$d_n\colon (M_f)_n \to (M_f)_{n-1}$$

defined by

$$d_n(c''_{n-1}, 0) = (-d''_{n-1}c''_{n-1}, f_{n-1}c''_{n-1})$$
$$d_n(0, c'_n) = (0, d_n c'_n).$$

$\mathcal{C}'$ is a subcomplex of M_f in a natural way with quotient complex $\Sigma'\mathcal{C}''$. The properties of the mapping cylinder are summarized in

LEMMA 2.6. *The sequence*

$$0 \to \mathcal{C}' \xrightarrow{i} M_f \xrightarrow{\pi} \Sigma'\mathcal{C}'' \to 0$$

is exact where i is the inclusion and π the projection. There is a commutative diagram

$$\begin{tikzcd}
H_{n+1}(\Sigma'\mathcal{C}'') \arrow[r, "\partial"] \arrow[d, equal] & H_n(\mathcal{C}') \\
H_n(\mathcal{C}'') \arrow[ur, "f_*"] &
\end{tikzcd}$$

where ∂ is the boundary in the homology sequence and f_* is the map on homology inclused by f.

THEOREM 2.7. *Suppose $f\colon \mathcal{C}'' \to \mathcal{C}'$ is a chain equivalence of free R complexes of finite type. If $\mathcal{C}'$ satisfies $G_{N'}$ and $\mathcal{C}''$ satisfies $G_{N''}$, then $\kappa_N(\mathcal{C}'') = \kappa_N(\mathcal{C}')$ where $N = \max(N', N'')$.*

Proof. Construct the mapping cylinder M_f. Since $f_*\colon H_n(\mathcal{C}'') \to H_n(\mathcal{C}')$ is an isomorphism, $\partial\colon H_{n+1}(\Sigma'\mathcal{C}'') \to H_n(\mathcal{C}')$ is an isomorphism and hence $H_n(M_f) = 0$ for all n. In particular $H_0(M_f) = 0$, so $B_0(M_f) = (M_f)_0$ and $(M_f)_0/B_0(M_f) = 0$. Thus M_f satisfies G_0 and $\kappa_0(M_f) = 0$.

Since $\mathcal{C}''$ satisfies $G_{N''}$, $\Sigma'\mathcal{C}''$ satisfies $G_{N''+1}$. Thus Theorem 2.3 applies and

$$\begin{aligned} 0 &= \kappa_{N+1}(\mathcal{C}') + \kappa_{N+1}(\Sigma'\mathcal{C}'') \\ &= -\kappa_N(\mathcal{C}') + \kappa_N(\mathcal{C}''). \end{aligned}$$

This completes the proof.

Corollary 2.8. *If $\mathcal{C}$ satisfies G_N and is chain equivalent to a finite free complex $\mathcal{C}'$ of dimension N', then $\kappa_N(\mathcal{C}) = 0$.*

Proof. Obviously $\mathcal{C}'$ satisfies $G_{N'}$ and $\kappa_{N'}(\mathcal{C}') = 0$. Thus

$$\begin{aligned} \kappa_N(\mathcal{C}) &= (-1)^{N'}\kappa_{N+N'}(\mathcal{C}) \\ &= (-1)^{N'}\kappa_{N+N'}(\mathcal{C}') = 0. \end{aligned}$$

Remark. 1.4 and 2.8 imply that Wall's obstruction is an invariant of homotopy type and is actually an obstruction to finiteness.

3. Suppose Π and Π' are groups and let $\mathcal{C}$ be a free $\mathbf{Z}\Pi$ complex of finite type and $\mathcal{C}'$ be a finite free $\mathbf{Z}\Pi'$ complex of dimension r. Thus α_i = rank of C'_i for each i is finite and $\chi(Y) = \Sigma(-1)^i\alpha_i$ is the Euler-Poincaré characteristic of Y.

Definition 3.1. $\mathcal{C}\otimes\mathcal{C}'$ is the complex with $(\mathcal{C}\otimes\mathcal{C}')_n = \sum_{i+j=n} C_i \underset{Z}{\otimes} C'_j$,

$$d_n\colon (\mathcal{C}\otimes\mathcal{C}')_n \to (\mathcal{C}\otimes\mathcal{C}')_{n-1}$$

defined by

$$d_n(c_i \underset{Z}{\otimes} c'_j) = d_i c_i \underset{Z}{\otimes} c'_j + (-1)^i c_i \underset{Z}{\otimes} d_j c'_j.$$

Each $(\mathcal{C}\otimes\mathcal{C}')_n$ becomes a $\mathbf{Z}[\Pi\times\Pi']$ module by

$$(\pi\times\pi')(c_i \underset{Z}{\otimes} c'_j) = \pi c_i \underset{Z}{\otimes} \pi' c'_j.$$

It is easy to see that each $(\mathcal{C}\otimes\mathcal{C}')_n$ is a finitely generated free $\mathbf{Z}[\Pi\times\Pi']$ module. For convenience let $\Lambda = \mathbf{Z}[\Pi\times\Pi']$.

Lemma 3.2. *Suppose $\mathcal{C}'$ is such that $C'_i = 0$, $i \neq r$; C'_r is a free $\mathbf{Z}\Pi'$ module of rank α_r. If $\mathcal{C}$ is a free $\mathbf{Z}\Pi$ complex of finite type satisfying G_N, then $\mathcal{C}\otimes\mathcal{C}'$ satisfies G_{N+r} over Λ and*

$$\kappa_{N+r}(\mathcal{C}\otimes\mathcal{C}') = \alpha_r i_* \kappa_N(\mathcal{C})$$

where $i_*\colon K_0(\mathbf{Z}\Pi) \to K_0(\Lambda)$ is induced by the inclusion

$$\pi \xrightarrow{\;i\;} (\pi, 1) \in \Pi\times\Pi', \quad \pi\in\Pi.$$

Proof. Consider first the special use $\alpha_r = 1$. Since $\mathcal{C}$ satisfies G_N and $C_N/B_N = D_N$ is projective, we may consider $C_N = B_N \oplus D_N$ as a $\mathbf{Z}\Pi$ module. Define a chain map $f\colon \mathcal{C} \to \mathcal{C}$ by

$$\begin{aligned} &f_{N+i} = 0, \; i > 0 \\ &f_n = 1, \; n < N \\ &f_N \mid B_N = 0, \; f_N \mid D_N = \text{identity}. \end{aligned}$$

Since $H_{N+i}(\mathcal{C}) = 0$, $i > 0$, one sees that $1 - f \cong 0$. Thus, there exist $\mathbf{Z}\Pi$ homomorphisms $\delta\colon C_j \to C_{j+1}$ such that

$$1_n - f_n = d_{n+1}\delta_n + \delta_{n-1}d_n$$

Thus

$$1_{N+i} = d_{N+i+1}\delta_{N+i} + \delta_{N+i-1}d_{N+i}, \; i > 0.$$

If we consider the family of maps, $\{\delta_{N+i} \otimes 1_r, i > 0\}$, we see that

$$Z_{N+i+r}(\mathcal{C} \otimes \mathcal{C}') \subset B_{N+i+r}(\mathcal{C} \otimes \mathcal{C}')$$

so

$$H_{N+i+r}(\mathcal{C} \otimes \mathcal{C}') = 0, \; i > 0.$$

Clearly $B_{N+r}(\mathcal{C} \otimes \mathcal{C}') \approx \Lambda \underset{\mathbf{Z}\Pi}{\otimes} B_N$ and $D_{N+r}(\mathcal{C} \otimes \mathcal{C}') \cong \Lambda \underset{\mathbf{Z}\Pi}{\otimes} D_N$; these facts imply that $\mathcal{C} \otimes \mathcal{C}'$ satisfies G_{N+r} and

$$\kappa_{N+r}(\mathcal{C} \otimes \mathcal{C}') = i_* \kappa_N(\mathcal{C}),$$

which completes the proof in the special use. The general case follows by repeated application of Theorem 2.3.

THEOREM 3.3. *Let $\mathcal{C}$ be a free $\mathbf{Z}\Pi$ complex of finite type satisfying G_N for some N and let $\mathcal{C}'$ be a finite free $\mathbf{Z}\Pi'$ complex of dimension r. Then $\mathcal{C} \otimes \mathcal{C}'$ satisfies G_{N+r} and*

$$\begin{aligned} \kappa_{N+r}(\mathcal{C} \otimes \mathcal{C}') &= (-1)^r \chi(\mathcal{C}') i_* \kappa_N(\mathcal{C}) \\ &= \chi(\mathcal{C}') i_* \kappa_{N+r}(\mathcal{C}). \end{aligned}$$

Proof. Induction on the dimension r of $\mathcal{C}'$. If $r = 0$, then Lemma 3.2 applies and the theorem is true. Suppose the theorem is true for complexes $\mathcal{C}'$ of dimensions $\leqq r-1$. If $\mathcal{C}'^{(r-1)}$ is the $(r-1)$ skeleton of $\mathcal{C}'$, one has an exact sequence of $\mathbf{Z}\Pi'$ complexes

$$0 \to \mathcal{C}'^{(r-1)} \to \mathcal{C}' \to \mathcal{C}'/\mathcal{C}'^{(r-1)} \to 0$$

whence one obtains a short exact sequence of Λ complexes

$$0 \to \mathscr{C} \otimes \mathscr{C}'^{(r-1)} \to \mathscr{C} \otimes \mathscr{C}' \to \mathscr{C}' \otimes (\mathscr{C}'/\mathscr{C}'^{(r-1)}) \to 0$$

By the induction hypothesis, $\mathscr{C} \otimes \mathscr{C}'^{(r-1)}$ satisfies G_{N+r-1} and

$$\kappa_{N+r-1}(\mathscr{C} \otimes \mathscr{C}'^{(r-1)}) = (-1)^{r-1}\chi(\mathscr{C}'^{(r-1)}) i_*\kappa_N(\mathscr{C})$$

By Lemma 3.2, $\mathscr{C} \otimes (\mathscr{C}'/\mathscr{C}'^{(r-1)})$ satisfies G_{N+r} and

$$\kappa_{N+r}(\mathscr{C} \otimes (\mathscr{C}'/\mathscr{C}'^{(r-1)})) = \alpha_r i_*\kappa_N(\mathscr{C}).$$

Then Theorem 2.3 implies that $\mathscr{C} \otimes \mathscr{C}'$ satisfies G_{N+r} and

$$\begin{aligned}\kappa_{N+r}(\mathscr{C} \otimes \mathscr{C}') &= \kappa_{N+r}(\mathscr{C} \otimes \mathscr{C}'^{(r-1)}) + \kappa_{N+r}(\mathscr{C} \otimes (\mathscr{C}'/\mathscr{C}'^{(r-1)}))\\ &= -\kappa_{N+r-1}(\mathscr{C} \otimes \mathscr{C}'^{(r-1)}) + \alpha_r i_*\kappa_N(\mathscr{C})\\ &= (-1)^r\chi(\mathscr{C}'^{(r-1)}) i_*\kappa_N(\mathscr{C}) + \alpha_r i_*\kappa_N(\mathscr{C})\\ &= (-1)^r[\chi(\mathscr{C}'^{(r-1)}) + (-1)^r\alpha_r] i_*\kappa_N(\mathscr{C})\\ &= (-1)^r\chi(\mathscr{C}') i_*\kappa_N(\mathscr{C}).\end{aligned}$$

4. In order to apply 3.3 let X be a CW complex of finite type satisfying i) and ii) of § 0 for some $N > 2$. By our remarks at the end of § 1, $C_*(\tilde{X})$ viewed as a $\pi_1(X)$ complex satisfies G_N. If Y is a finite complex of dimension r, then

$$C_*(\widetilde{X \times Y}) = C_*(\tilde{X} \times \tilde{Y}) = C_*(\tilde{X}) \otimes C_*(\tilde{Y}).$$

as $\mathbf{Z}\pi_1(X \times Y) = \mathbf{Z}(\pi_1(X) \times \pi_1(Y)) = \Lambda$ complexes (we have suppressed the base points in the notation but they could be kept in mind). By 3.3 $C_*(\widetilde{X \times Y})$ satisfies G_{N+r} and

4.1 $$\kappa_{N+r}(C_*(\widetilde{X \times Y})) = (-1)^r\chi(Y) i_*\kappa_N(C_*(\tilde{X})$$

where $i\colon \pi_1(X) \to \Lambda$ is induced by the inclusion $X \to X \times y_0$, y_0 the base point of Y.

If $X \times Y$ satisfies i) and ii) of § 0, our class $\kappa_{N+r}(C_*(\widetilde{X \times Y}))$ agrees with Wall's obstruction $\kappa_{N+r}(X \times Y)$. In any case, $\kappa_{N+r}(C_*(\widetilde{X \times Y})) = 0$ if and only if $X \times Y$ has the homotopy type of a finite CW complex.

If X is dominated by a finite CW complex K and Y is finite, then $X \times Y$ is dominated by $K \times Y$, so both X and $X \times Y$ satisfy i) and i)) of § 0. An interesting problem is to find a Künneth formula for local coefficient which would imply that i) and ii) are satisfied for $X \times Y$ if they are for X and Y is finite, without the assumption of domination by finite complexes.

If we interpret equation 4.1 under the hypotheses of 0.1, we see from $\kappa_N(X) = \kappa_N(C_*(\tilde{X}))$ and

$$\kappa_{N+r}(X \times Y) = \kappa_{N+r}(C_*(\widetilde{X \times Y}))$$

that the conclusion of 0.1 holds.

5. Theorem 3.3 is closely related to a product formula for Reidemeister torsion first stated by Milnor ([3], Theorem B) and a product formula for Whitehead torsion [2], a proof of which can be given analogous to the proof of 3.3. It would be interesting to know the connections between the Whitehead torsion group and the projective class group (1).

REFERENCES.

[1] H. Bass, "K-theory and stable algebra," *Institut des Hautes Études Scientifiques, Publications Mathématiques*, 1964.

[2] K. W. Kwun and R. H. Szczarba, "Product and sum theorems for Whitehead torsion" (to appear).

[3] J. W. Milnor, "Two complexes which are homeomorphic but combinatorially distinct," *Annals of Mathematics*, vol. 74 (1961), pp. 575-590.

[4] C. T. C. Wall, "An obstruction to finiteness of *CW* complexes," *Bulletin of the American Mathematical Society*, vol. 70 (1964), pp. 269-270.

EMBEDDINGS OF TOPOLOGICAL PRODUCTS OF K-CELL-LIKE CONTINUA.

By L. FEARNLEY.*

---

1. **Introduction.** The purpose of this paper is to investigate the Euclidean spaces of least dimension in which topological products of finite collections of k-cell-like continua can be embedded. In addition, for the special case where $k = 2$, we establish particular conditions under which disk-like continua can be embedded in Euclidean space of dimension three. Specifically, the following two principal theorems will be proved.

THEOREM 1. *The topological product of any collection of n k-cell-like continua can be embedded in Euclidean space of dimension $k(n+1)$.*

THEOREM 2. *A disk-like continuum L can be embedded in Euclidean space of dimension three if L is homeomorphic with the inverse limit of a system $\{D_i, f_i\}$ in which each space D_i is the unit disk in E^2 and each bonding mapping f_i can be factored into the composition of a homeomorphism into E^3 and a vertical projection of E^3 onto E^2.*

The first of these theorems gives an affirmative answer to a conjecture raised by R. H. Bing in a research seminar at the University of Wisconsin, Madison, Wisconsin, 1964. This theorem is also a generalization of a result recently obtained by R. Bennett [1]. In particular, in the case that $k = 1$, Theorem 1 reduces to the result established by Bennett that the topological product of any collection of n snake-like continua can be embedded in Euclidean space of dimension $n+1$. For the special case that $n = 1$, Theorem 1 gives the same result as one obtained by Isbell in [9] where it is shown that the inverse limit of an inverse system in which each co-ordinate space is a subset of Euclidean space of dimension n can be embedded in Euclidean space of dimension $2n$. For $k = 1$ and all values of n, Theorem 1

Received March 10, 1965.

* This research was supported in part by the National Science Foundation under Grant NSF-GP 2244.

is the best result [1] and this development suggests that for no values k and n can this be improved. However, we do not prove this last conjecture.

The embedding problem for disk-like continua is completely solved by Theorem 1. We note that the cone over a solenoid, obtained by taking the inverse limit of an inverse system in which each co-ordinate space is the unit disk in E^2 and each bonding mapping is the function defined by the complex number equation $w = z^2$, is an example of a disk-like continuum which cannot be embedded in E^3, so that the result for disk-like continua given by Theorem 1 cannot be improved. However, it is of interest to determine under what conditions a disk-like continuum can be embedded in E^3. In Theorem 2, we give an affirmative answer to a second conjecture of Bing which establishes such a condition.

It will develop that the embedding condition described in Theorem 2 is analogous in certain respects to the characterization of circularly chainable continua which can be embedded in the plane established by Bing in [3]. The results of this paper are also related to investigations of the mappings and topological products of snake-like continua and generalized snake-like continua undertaken by the author in [6] and [7].

The more standard terms used in this paper are defined in [5] or [11]. In addition, we give the following definitions of terms which have either not yet been completely standardized or which are frequently given different although equivalent definitions in the literature. A continuum is said to be *k-cell-like* if it is homeomorphic with the inverse limit space of a sequence $\{S_i, f_i\}$ in which each space S_i is a k-cell. We define a homeomorphism h of a compact subset K of a metric space S onto a subset $h(K)$ of S to be an *extensible homeomorphism* with respect to S if there is an open set U in S containing K and an extension of h to a homeomorphism of U into S. In referring to the distance between functions it will be assumed that the functions have a common domain and that the usual function space metric applies.

2. Preliminary results.

Before proceeding to give proofs of the principal theorems of this paper a number of preliminary lemmas will be needed. The lemmas of this section will have a general character and will be of some interest in themselves. First, we give a statement of a result of McCord [10].

LEMMA 1. *Let S be a compact metric space, let $\{S_i, f_i\}$ be an inverse system each space of which is a compact subset of S, and suppose, for each positive number ϵ and each positive integer i, that there is an extensible homeomorphism $h_{\epsilon i}$ with respect to S such that f_i and $h_{\epsilon i}$ have distance apart less than ϵ. Then the inverse limit of $\{S_i, f_i\}$ can be embedded in S.*

In the following lemma we establish a result which will be of particular use in proving the first principal theorem of the paper.

LEMMA 2. *If* $\{S_i, f_i\}$ *is an inverse system each space of which is the unit n-cube in* E^n *and* k *is an integer,* $1 \leqq k \leqq n$, *such that each bonding mapping of the system is a function of* k *variables, then the inverse limit of* $\{S_i, f_i\}$ *can be embedded in* E^{n+k}.

Proof. Assume that a particular cartesian co-ordinate system has been chosen for E^n and let i be a positive integer. We shall show that f_i can be approximated arbitrarily closely by homeomorphisms which are extensible with respect to a compact subset of E^{n+k}. To do this we first choose a family of mappings $\{g_\epsilon\colon 0 < \epsilon < 1\}$ in the following manner. For each value of ϵ, g_ϵ is a transformation of the open n-cube

$$\{(x_1, x_2, \cdots, x_n) : -\epsilon < x_i < 1 + \epsilon; i = 1, 2, \cdots, n\}$$

onto the open unit n-cube $\{(x_1, x_2, \cdots, x_n) : 0 < x_i < 1; i = 1, 2, \cdots, n\}$ defined by the equation

$$\begin{aligned} &g_\epsilon(x_1, x_2, \cdots, x_n) \\ &\quad = ((x_1 + \epsilon)/(1 + 2\epsilon), (x_2 + \epsilon)/(1 + 2\epsilon), \cdots, (x_n + \epsilon)/(1 + 2\epsilon)). \end{aligned}$$

Next, we consider the unit n-cube to be represented in terms of co-ordinates of E^{n+k} so that f_i can be written in the form

$$\begin{aligned} f_i(x_1, x_2, \cdots, x_{n+k}) = (f_{i1}(x_1, x_2, \cdots, x_n), f_{i2}(x_1, x_2, \cdots, x_n), \cdots,& \\ f_{in}(x_1, x_2, \cdots, x_n), 0, 0, \cdots, 0).& \end{aligned}$$

Now, let $u_1, u_2, \cdots, u_k$ be k integers, $1 \leqq u_j \leqq n$, $j = 1, 2, \cdots, k$, such that f_i is a function of the variables $x_{u_1}, x_{u_2}, \cdots, x_{u_k}$, let $v_1, v_2, \cdots, v_{n-k}$ be the integers of the set $\{1, 2, \cdots, n\} - \{u_1, u_2, \cdots, u_k\}$ and let ϵ be a positive number. We proceed to define a homeomorphism $h_{\epsilon i}$ whose domain is the open $(n + k)$-cube $\{(x_1, x_2, \cdots, x_{n+k}) : -\epsilon < x_i < 1 + \epsilon; i = 1, 2, \cdots, n + k\}$ and show that the family of homeomorphisms obtained by restricting such mappings $\{h_{\epsilon i}\colon 0 < \epsilon < 1\}$ to the domain $\{(x_1, x_2, \cdots, x_{n+k}) : 0 \leqq x_i \leqq 1;$ $i = 1, 2, \cdots, n; x_{n+1} = 0, x_{n+2} = 0, \cdots, x_{n+k} = 0\}$ includes homeomorphisms arbitrarily close to f_i. Specifically, we define $h_{\epsilon i}$ in the following manner.

$$\begin{aligned} h_{\epsilon i}(x_1, x_2, \cdots, x_{n+k}) &= (f_{i1} g_\epsilon(x_1, x_2, \cdots, x_n) + \epsilon x_{v_1}, f_{i2} g_\epsilon(x_1, x_2, \cdots, x_n) \\ &\quad + \epsilon x_{v_2}, \cdots, f_{i\,n-k} g_\epsilon(x_1, x_2, \cdots, x_n) + \epsilon x_{v_{n-k}}, f_{i\,n-k+1} g_\epsilon(x_1, x_2, \cdots, x_n) \end{aligned}$$

$$+ \epsilon x_{n+1}, f_{i\,n-k+2}g_\epsilon(x_1, x_2, \cdots, x_n) + \epsilon x_{n+2}, \cdots, f_{in}g_\epsilon(x_1, x_2, \cdots, x_n)$$
$$+ \epsilon x_{n+k}, \epsilon x_{u_1}, \epsilon x_{u_2}, \cdots, \epsilon x_{u_k}).$$

Then, since the values of the first n co-ordinates of the points of the domain of $h_{\epsilon i}$ are values of the corresponding co-ordinates of the points of the domain of g_ϵ, and the range of g_ϵ is contained in the domain of f_i, it follows that $h_{\epsilon i}$ is a well defined transformation. Furthermore, $h_{\epsilon i}$ is a continuous transformation. To see that $h_{\epsilon i}$ is also a one-to-one transformation suppose that $(y_1, y_2, \cdots, y_{n+k})$ and $(z_1, z_2, \cdots, z_{n+k})$ are elements of the domain of $h_{\epsilon i}$ such that $h_{\epsilon i}(y_1, y_2, \cdots, y_{n+k}) = h_{\epsilon i}(z_1, z_2, \cdots, z_{n+k})$. Then, equating corresponding members of the last k co-ordinates of $h_{\epsilon i}(y_1, y_2, \cdots, y_{n+k})$ and $h_{\epsilon i}(z_1, z_2, \cdots, z_{n+k})$, we obtain relationships $y_{u_j} = z_{u_j}$, $j = 1, 2, \cdots, k$. Hence $g_\epsilon(y_1, y_2, \cdots, y_n)$ and $g_\epsilon(z_1, z_2, \cdots, z_n)$ have equal values in each of the co-ordinate positions u_i, $i = 1, 2, \cdots, k$. In addition, f_i is a function of the variables $x_{u_1}, x_{u_2}, \cdots, x_{u_k}$ so that $f_{is}g_\epsilon(y_1, y_2, \cdots, y_n) = f_{is}g_\epsilon(z_1, z_2, \cdots, z_n)$ for $s = 1, 2, \cdots, n$. It follows, by equating corresponding members of the co-ordinates of $h_{\epsilon i}(y_1, y_2, \cdots, y_{n+k})$ and $h_{\epsilon i}(z_1, z_2, \cdots, z_{n+k})$, that $y_t = z_t$ for $t = 1, 2, \cdots, n + k$. Therefore, the restriction $r_{\epsilon i}$ of $h_{\epsilon i}$ to the unit n-cube

$$\{(x_1, x_2, \cdots, x_{n+k}) : 0 \leqq x_s \leqq 1, s = 1, 2, \cdots, n,$$
$$x_{n+1} = 0, x_{n+2} = 0, \cdots, x_{n+k} = 0\}$$

is an extensible homeomorphism with respect to E^{n+k}. Furthermore, since the domain and range of each of the homeomorphisms $\{h_{\epsilon i}: 0 < \epsilon < 1; i = 1, 2, 3, \cdots\}$ is contained in a $3(n + k)$ neighborhood of the origin in E^{n+k}, the homeomorphisms $\{r_{\epsilon i}: 0 < \epsilon < 1; i = 1, 2, 3, \cdots\}$ are extensible with respect to a compact subset M of E^{n+k}.

Now, as ϵ approaches zero, the functions $\{g_\epsilon: 0 < \epsilon < 1\}$ converge uniformly to the identity mapping. Furthermore, each of the functions f_{is}, $s = 1, 2, \cdots, n$, is uniformly continuous. Thus, for each positive number δ, there is an extensible homeomorphism with respect to M in the collection $\{r_{\epsilon i}\}$ having distance less than δ from f_i. We conclude, noting Lemma 1, that the inverse limit of $\{S_i, f_i\}$ can be embedded in M and hence in E^{n+k}.

The final result of this section involves embeddings of inverse limits in which the bonding mappings are generalizations of the functions considered in Lemma 2. In particular, we define a mapping f_i of a subset of E^n into E^n to be *effectively a function of k variables*, $1 \leqq k \leqq n$, if there exist integers $u_1, u_2, \cdots, u_k$ of the set $\{1, 2, \cdots, n\}$ such that the defining equation of f_i,

$$f_i(x_1, x_2, \cdots, x_n) = (f_{i1}(x_1, x_2, \cdots, x_n), f_{i2}(x_1, x_2, \cdots, x_n), \cdots,$$
$$f_{in}(x_1, x_2, \cdots, x_n))$$

has the following properties:

(1) Each of the real valued functions f_{iu_r}, $r=1,2,\cdots,k$, is a function of k variables $x_{u_1}, x_{u_2},\cdots,x_{u_k}$.

(2) If v_s is an integer of the set $\{1,2,\cdots,n\}-\{u_1,u_2,\cdots,u_k\}$, then $f_{iv_s}(x_1,x_2,\cdots,x_n)=x_{v_s}$.

LEMMA 3. *If each bonding mapping of an inverse system $\{S_i,f_i\}$ is effectively a function of k variables which takes the unit n-cube onto itself, $1\leqq k\leqq n$, then the inverse limit of $\{S_i,f_i\}$ can be embedded in E^{n+k}.*

Proof. Let f_i be a bonding mapping of $\{S_i,f_i\}$ with properties described in the same notation as that of the above definition. We then use the same argument as that given in the proof of Lemma 2 with the exception that $h_{\epsilon i}$ is chosen to be a transformation such that the co-ordinate of $h_{\epsilon i}(x_1,x_2,\cdots,x_{n+k})$ in position u_r is $f_{iu_r}g_\epsilon(x_1,x_2,\cdots,x_n)+\epsilon x_{n+r}$, $r=1,2,\cdots,k$, the co-ordinate of $h_{\epsilon i}(x_1,x_2,\cdots,x_{n+k})$ in position v_s is x_{v_s}, $s=1,2,\cdots,n-k$, and the co-ordinate of $h_{\epsilon i}(x_1,x_2,\cdots,x_{n+k})$ in position $n+r$ is ϵx_{u_r}, $r=1,2,\cdots,k$.

3. Topological products of k-cell-like continua. In this section we use the results of the preceding section in establishing the first principal theorem mentioned in the Introduction of this paper. The spaces of the inverse systems considered in this theorem will be unit k-cells and unit nk-cells. However, for the purposes of describing the transformations developed in the proof of the theorem, it will be convenient to index otherwise identical spaces according to their positions in a given inverse limit sequence.

THEOREM 1. *If $K_1,K_2,\cdots,K_n$ is a collection of k-cell-like continua, then the topological product $K_1\times K_2\times\cdots\times K_n$ can be embedded in a Euclidean space of dimension $k(n+1)$.*

Proof. We may assume that each k-cell-like continuum K_t, $t=1,2,\cdots,n$, is the inverse limit space of a sequence $\{C_{st},f_{st}\}$, $s=1,2,3,\cdots$, in which each space C_{st} is a unit k-cell and each bonding function f_{st} is a mapping of $C_{s+1\,t}$ onto C_{st}. With these inverse limit sequences we construct a new inverse limit sequence involving nk-cells which will be the principal component in establishing the desired embedding result.

First we define nk-cells $\{D_{pq}\colon p=1,2,3,\cdots;q=1,2,\cdots,n\}$ in the following manner

$$D_{pq} = C_{p+1\,1} \times C_{p+1\,2} \times \cdots \times C_{p+1\,q-1} \times C_{pq} \times C_{p\,q+1} \times \cdots \times C_{pn}$$

for $p = 1, 2, 3, \cdots ; q = 1, 2, \cdots , n$.

In the case that $q = 1$ the defining equation for D_{pq} is to be interpreted as indicating that $D_{pq} = C_{pq} \times C_{p\,q+1} \times \cdots \times C_{pn}$. Next, these nk-cells $\{D_{pq}\}$ are ordered lexicographically with respect to their subscripts. To complete the defintion of this inverse limit sequence, for each subscript pair pq in the lexicographic ordering we denote by x_{pq} a representative element of C_{pq} and define the bonding mapping g_{pq} of the successor of D_{pq} onto D_{pq} by the equation

$$\begin{aligned} g_{pq}(x_{p+1\,1}, x_{p+1\,2}, \cdots, x_{p+1\,q}, x_{p\,q+1}, x_{p\,q+2}, \cdots, x_{pn}) \\ = (x_{p+1\,1}, x_{p+1\,2}, \cdots, f_{pq}(x_{p+1\,q}), x_{p\,q+1}, x_{p\,q+2}, \cdots, x_{pn}). \end{aligned}$$

The first entry on the right hand side is to be $f_{pq}(x_{p+1\,q})$ if $q = 1$. We denote the resulting lexicographically ordered inverse limit sequence by $\{D_{pq}, g_{pq}\}$ and indicate the inverse limit space of this sequence by L. The theorem will be established by showing that (1) L is homeomorphic with the topological product $K_1 \times K_2 \times \cdots \times K_n$, and (2) L can be embedded in $E^{k(n+1)}$.

To establish Assertion (1) we note first, since L is the inverse limit space of $\{D_{pq}, g_{pq}\}$, that each element of L has the form

$$\begin{aligned} a = ((x_{11}, x_{12}, \cdots, x_{1n}), (x_{21}, x_{12}, x_{13}, \cdots, x_{1n}), \cdots, \\ (x_{p+1\,1}, x_{p+1\,2}, \cdots, x_{p+1\,q-1}, x_{pq}, x_{p\,q+1}, \cdots, x_{pn}), \\ (x_{p+1\,1}, x_{p+1\,2}, \cdots, x_{p+1\,q}, x_{p\,q+1}, x_{p\,q+2}, \cdots, x_{pn}), \\ \cdots). \end{aligned}$$

We shall refer to the expressions enclosed by inner parenthesis in this equation for a as "co-ordinate terms" of a and use the projection mapping notation to indicate the position of a particular co-ordinate term in the sequence of co-ordinate terms. Thus $\pi_1(a)$ would denote the first co-ordinate term $(x_{11}, x_{12}, \cdots, x_{1n})$. The individual entries, ignoring inner parentheses, of the expression on the right hand side of the equation for a will be referred to as "elementary terms." Thus x_{11} would be the first elementary term of a.

Now, in the expression for a, $f_{pq}(x_{p+1\,q}) = x_{pq}$ for $p = 1, 2, 3, \cdots ;$ $q = 1, 2, \cdots, n$. Hence the followling expression

$$b = ((x_{11}, x_{21}, x_{31}, \cdots), (x_{12}, x_{22}, x_{32}, \cdots), \cdots, (x_{1n}, x_{2n}, x_{3n}, \cdots))$$

represents a point of the topological product $K_1 \times K_2 \times \cdots \times K_n$. We define a transformation h of L onto $K_1 \times K_2 \times \cdots \times K_n$ by setting

$$h(a) = b.$$

It is easily verified that h is a one-to-one transformation of L onto $K_1 \times K_2 \times \cdots \times K_n$. To see that h is also a continuous transformation we note, from [9, Th. 14, pp. 122-123], that if $u = (u_1, u_2, u_3, \cdots)$ and $v = (v_1, v_2, v_3, \cdots)$ are elements of the topological product space of a uniformly bounded sequence of metric spaces with metric functions $d_1, d_2, d_3, \cdots$, respectively, then the topology induced by the metric $d(u,v) = \sum_1^\infty 2^{-i} d_i(u_i, v_i)$ is equivalent to the usual Tychonoff topology. Thus, since the co-ordinate spaces $\{D_{pq}\}$ of L are uniformly bounded metric spaces, the above type of metric function may be used for L. Hence, if $a_1, a_2, a_3, \cdots$ is a sequence of points of L converging to a point a of L and i is a positive integer, then $\pi_i(a_1), \pi_i(a_2), \pi_i(a_3), \cdots$ converges to $\pi_i(a)$. Furthermore, if $(x_{pq})_i$ is an elementary term of a_i, $i = 1, 2, 3, \cdots$, then $(x_{pq})_1, (x_{pq})_2, (x_{pq})_3, \cdots$ converges to the elementary term x_{pq} of a. It follows that $h(a_1), h(a_2), h(a_3), \cdots$ converges to $h(a)$. Thus h is a continuous transformation and we conclude that the inverse limit space L is homeomorphic with the topological product $K_1 \times K_2 \times \cdots \times K_n$.

To establish Assertion (2), we observe that each space of the inverse limit system $\{D_{pq}, g_{pq}\}$ is a unit nk-cell. Furthermore, each bonding mapping g_{pq} of this system is a function of the representative $x_{p+1\,q}$ of the k-cell $C_{p+1\,q}$. Thus, each bonding mapping of the system $\{D_{pq}, g_{pq}\}$ is effectively a function of k variables. Therefore, by Lemma 3, the inverse limit L can be embedded in a Euclidean space of dimension $nk + k = k(n+1)$. This completes the proof.

4. Embeddings of disk-like continua in E^3. The purpose of this section is to establish a condition under which disk-like continua can be embedded in E^3. We shall consider bonding mappings of a canonical planar disk onto itself which can be factored into the composition of a homeomorphism into E^3 and a vertical projection of E^3 onto E^2. It is observed that if such a mapping f takes a simple closed curve J onto a simple closed curve $f(J)$, then the restriction of f to J can be approximated by a homeomorphism h into a toroidal neighborhood of $h(J)$. In this situation, a circling number of f on J can be assigned, equal to the circling number of $h(J)$ in the associated neighborhood of $f(J)$. Then, since f can be factored in the manner indicated above, it is easily verified that the circling number of f on J is at most one. In this respect, the condition described in Theorem 2 is similar to the characterization of the circularly chainable continua which can be embedded in the plane established by Bing [3].

The proof of Theorem 2 will depend primarily on the Improved Side Approximation Theorem of Bing [2].

THEOREM 2. *A disk-like continuum L can be embedded in Euclidean space of dimension three if L is homeomorphic with the inverse limit of a system $\{D_i, f_i\}$ in which each space D_i is the unit 2-cell in E^2 and each bonding mapping f_i can be factored into the composition of a homeomorphism into E^3 and a vertical projection of E^3 onto E^2.*

Proof. First, we choose a positive number ϵ and a positive integer i. In terms of the standard cartesian co-ordinate representation of E^3, there is a homeomorphism of E^3 onto the set $\{(x, y, z) : -\epsilon/3 < z < \epsilon/3)$ which is the identity mapping on the set $\{(x, y, z) : z = 0\}$ and which moves points only in a vertical manner. Thus, we may assume without loss in generality that $f_i = p_i h_i$, where h_i is a homeomorphism of the disk D_{i+1} into E^3 and p_i is a vertical projection of $h_i(D_{i+1})$ onto D_i which moves no point more than $\epsilon/3$.

Now, by the Improved Side Approximation Theorem of Bing [2], there exists a homeomorphism g_i of the interior of D_{i+1} into E^3 such that g_i has distance less than $\epsilon/3$ from the restriction of h_i to $\operatorname{Int} D_{i+1}$ and $g_i(\operatorname{Int} D_{i+1})$ is a tame disk without boundary. We let δ be a positive number such that each subset of D_{i+1} of diameter less than δ is mapped by f_i into a subset of D_i of diameter less than $\epsilon/3$. This is possible since f_i is a uniformly continuous function. Next we define c_i to be a homeomorphism of the open disk $E_{i+1} = \{(x, y, z) : -\delta < x < 1 + \delta, -\delta < y < 1 + \delta, z = 0\}$ onto the interior of the disk D_{i+1} so that c_i moves no point more than δ. We shall show that the restriction of the composite function $g_i c_i$ to the disk D_{i+1} is an extensible homeomorphism with respect to E^3 having distance less than ϵ from f_i.

To do this, observe first, from the fact that E_{i+1}, $c_i(E_{i+1})$ and $g_i c_i(E_{i+1})$ are two sided tame 2-manifolds in E^3 and the result [4, Th. 3, p. 339] of Brown, that each of these open disks can be bi-collared in E^3. Let U_{i+1}, V_{i+1} and W_{i+1} be open sets in E^3 obtained as a result of bi-collaring the open disks E_{i+1}, $c_i(E_{i+1})$ and $g_i c_i(E_{i+1})$, respectively. Then the homeomorphism c_i can be extended in a natural manner to a homeomorphism C_i of U_{i+1} onto V_{i+1}, and the homeomorphism g_i can be extended in a natural manner to a homeomorphism G_i of V_{i+1} onto W_{i+1}. Therefore, since U_{i+1} is an open set in E^3 containing the disk D_{i+1} and G_iC_i is a homeomorphism which is an extension of g_ic_i and has domain U_{i+1}, it follows that the restriction of g_ic_i to D_{i+1} is an extensible homeomorphism with respect to E^3.

Now, since c_i moves no point more than δ, the distance from f_i to the restriction of f_ic_i to D_{i+1} is less than $\epsilon/3$. Next, the distance from g_ic_i to f_ic_i is equal to the distance from g_i to the restriction of f_i to $\operatorname{Int} D_{i+1}$. Furthermore, on the interior of D_{i+1}, the distance from g_i to f_i is less than or equal to the sum of the distances from g_i to h_i and h_i to f_i. Each of the distances of this sum is less than $\epsilon/3$. Hence the restriction of g_ic_i to D_{i+1} has distance less than ϵ from f_i.

To complete the proof of this theorem we note that without loss in generality we may have assumed that U_{i+1} and W_{i+1} were chosen to be open subsets of the unit neighborhood $N(D_{i+1})$ of D_{i+1}. Therefore, by Lemma 1, we conclude that the disk-like continuum L can be embedded in the closure of $N(D_{i+1})$ and hence in E^3.

Brigham Young University
and
University of Wisconsin.

REFERENCES.

[1] R. Bennett, "Embedding products of chainable continua," *Proceedings of the American Mathematical Society*, in process of publication.

[2] R. H. Bing, "Improving the side approximation theorem," to appear in *Transactions of the American Mathematical Society*.

[3] ———, "Embedding snake-like continua in the plane," *Canadian Journal of Mathematics*, vol. 14 (1962), pp. 113-128.

[4] Morton Brown, "Locally flat imbeddings of topological manifolds," *Annals of Mathematics*, vol. 75 (1962), pp. 331-341.

[5] S. Eilenberg and N. Steenrod, *Foundations of Algebraic Topology*, Princeton University Press, Princeton, 1952.

[6] L. Fearnley, "Characterization of the continuous images of the pseudo-arc," *Transactions of the American Mathematical Society*, vol. 111 (1964), pp. 380-399.

[7] ———, "Topological operations on the class of continuous images of all snake-like continua," *Proceedings of the London Mathematical Society*, (3) 15 (1965), pp. 289-300.

[8] J. R. Isbell, "Embeddings of inverse limits," *Annals of Mathematics*, vol. 70 (1959), pp. 73-84.

[9] J. L. Kelley, *General Topology*, Van Nostrand, New York, 1955.

[10] M. C. McCord, "Inverse limit systems," Thesis, Yale University, 1963.

[11] G. T. Whyburn, *Analytic Topology*, American Mathematical Colloquium Publications, vol. 28, American Mathematical Society, Providence, R. I., 1942.

MAPPINGS OF MANIFOLDS INTO THE PLANE.

By HAROLD I. LEVINE.*

1. Introduction and statement of result. 1.1. Let M be a compact manifold,[1] and let f be a real valued function on M with only non-degenerate critical points. This means, if $P \in M$ is a critical point of f, that is $df(P) = 0$, then the Hessian of f at P, $Hf(P)$, is non-singular. The index of P, a critical point of f, is the index of $Hf(P)$. Let $N_i(f)$ be the number of critical points of f of index i. The Morse-Pitcher inequalities [5] relate the N_i's with topological invariants of M as follows:

$$(1) \qquad \sum_{i=0}^{k} (-1)^{k-i} N_i(f) \geqq \sum_{i=0}^{k} (-1)^{k-i} B_i + T_k,$$

where B_i is the i-th Betti number of M and T_k is the number of torsion coefficients in $H_k(M; \mathbf{Z})$. If $k = n = \dim M$, we obtain the equation:

$$(2) \qquad \sum_{i=0}^{n} (-1)^{i} N_i(f) = \mathrm{X}(M),$$

where $\mathrm{X}(M)$ is the Euler characteristic of M.

In this note we study maps of orientable manifolds into the plane and prove an analog of (2) for even-dimensional manifolds.

1.2. In order to obtain the results quoted above for maps into the line, we restrict ourselves to a dense set of maps, those with only non-degenerate critical points. Here also we take as objects of study, elements of a dense set, $\mathcal{F}(M)$ of maps of M into the plane. The set $\mathcal{F} = \mathcal{F}(M)$, is defined as follows:

A map $f: M \to \mathbf{R}^2$ is in $\mathcal{F}$ if and only if:

Received April 29, 1965.

* This work was supported in part by the National Science Foundation under Grant No. NSF GP-170.

[1] Except when they are explicitly described, all mappings and manifolds will be understood to be C^∞.

(i) The rank of df, $rk\, df \geqq 1$.

(ii) If $P \in S_1(f) = \{Q \in M \mid rk\, df(Q) = 1\}$, then we can choose coordinates $\{u, x, z_1, \cdots, z_{n-2}\}$ centered at P and $\{U, Y\}$ centered at $f(P)$ such that f has one of the forms:

$$\left\{\begin{aligned} U \circ f(u,x,z) &= u \\ Y \circ f(u,x,z) &= \sum_{i=1}^{n-2} \pm z_i^2 \pm x^2 \end{aligned}\right. \tag{3}$$

or

$$\left\{\begin{aligned} U \circ f(u,x,z) &= u \\ Y \circ f(u,x,z) &= \sum_{i=1}^{n-2} \pm z_i^2 + axu + bx^3, \end{aligned}\right. \tag{4}$$

where $|a| = |b| = 1$.

A more concise way of describing $\mathfrak{F}$ is that $f \in \mathfrak{F}$ if and only if $J^1(f)$ is transversal to $S_1(M, \boldsymbol{R}^2)$ and $S_2(M, \boldsymbol{R}^2)$ and that $J^2(f)$ is transversal to $S_1{}^2(M, \boldsymbol{R}^2)$. The equivalence of these descriptions of $\mathfrak{F}$ is just a theorem of Whitney [8] in case $\dim M = 2$ which gives the normal forms (3) and (4) for maps satisfying the "jet-transversality" conditions. The extension of this theorem to $\dim M > 2$ is easy [3]. That $\mathfrak{F}$ is dense in all maps of M into $\boldsymbol{R}^2$ is a special case of the transversality theorem of Thom [6, 2].

Let $f \in \mathfrak{F}$. A *regular point* of f is a point $P \in M$ such that $rk\, df(P) = 2$. All other points of M are called *singular*. In our case the singular set of f is the same as $S_1(f)$ and so $S_1(f)$ is closed in M. Since M is compact it is thus clear from the normal forms (3) and (4) or on general transversality grounds that $S_1(f)$ consists of a finite collection of closed, non-singular, non-intersecting curves. The set of points on $S_1(f)$ at which f has one of the normal forms (4) is discrete; this finite set is called $S_1{}^2(f)$. This set is the subset of $S_1(f)$ at which df annihilates the tangent space to $S_1(f)$. The components of $S_1(f)$ are called the *fold curves of f* and the points of $S_1{}^2(f)$ are called the *cusp points of f*.

1.3. Choose an orientation for $\boldsymbol{R}^2$. Relative to this orientation, take the unit circle in $\boldsymbol{R}^2$ with the counterclockwise orientation. Orient the space of lines through 0 in $\boldsymbol{R}^2$, $P(\boldsymbol{R}^2)$, so that the double covering of it by the unit circle (= space of directions) is orientation preserving.

For any $f \in \mathfrak{F}$, we define a map $\phi\colon S_1(f) \to P(\boldsymbol{R}^2)$, where if $P \in S_1(f)$

$- S_1{}^2(f)$, $\phi(P)$ is the line through 0 in $\boldsymbol{R}^2$, parallel to the tangent to $f(S_1(f))$ at $f(P)$.

Lemma. *As defined above, the map* $\phi\colon S_1(f) - S_1{}^2(f) \to P(\boldsymbol{R}^2)$ *has a unique* C^∞ *extension to all of* $S_1(f)$.

Proof. Let $P \in S_1{}^2(f)$, and pick coordinates centered at P and $f(P)$ so that f has one of the forms (4). We work entirely in these coordinate *nbhds.* $S_1(f)$ is defined by $z_i = 0$, $i = 1, \cdots, n-2$ and $au + 3bx^2 = 0$. Let $Q \in S_1(f)$, with (x,u,z)-coordinates, $(x, -\dfrac{3b}{a}x^2, 0)$. A tangent vector to $S_1(f)$ is $(\dfrac{\partial}{\partial x} - \dfrac{6b}{a}x\dfrac{\partial}{\partial u})$ and its df image at $f(Q)$ is $-\dfrac{6b}{a}x(\dfrac{\partial}{\partial U} + ax\dfrac{\partial}{\partial Y})$. Thus the tangent line to $f(S_1(f))$ at $f(Q)$ is the line of multiples of $(\dfrac{\partial}{\partial U} + ax\dfrac{\partial}{\partial Y})$. This line is well defined even at P where $x = 0$ and plays the role of the tangent line to $f(S_1(f))$ at $f(P)$.

Theorem 1. *Let M be an even-dimensional, compact, oriented manifold and let $f \in \mathfrak{F}(M)$. One can orient $S_1(f)$ in such a way that:*

$$\sum_C r(C) = \mathrm{X}(M),$$

where the summation runs over the components C of $S_1(f)$ and $r(C)$ is the degree of $\phi \mid C$, and $\mathrm{X}(M)$ is the Euler characteristic of M.

Corollary 1. (Thom [7]) *The number of cusps of $f \in \mathfrak{F}(M)$ has the same parity as* $\mathrm{X}(M)$.

The corollary is immediate from the theorem since $r(C)$ is even if and only if the number of cusps on C is even.

Given an $f \in \mathfrak{F}(M)$, let $\mathfrak{G}(f)$ be the collection of all linear functions on the plane such that $\gamma \in \mathfrak{G}(f)$ if and only if $\gamma \circ f$ has only non-degenerate critical points. To show that almost all linear functions are in $\mathfrak{G}(f)$, it suffices to apply the argument of Kuiper [1, p. 79] to the components of $S_1(f) - S_1{}^2(f)$ immersed in the plane (see 2.2 below). The proof of Theorem 1, which is given in the next section, is just a reinterpretation of the Morse equation (2), for $\gamma \circ f$, where γ is any element of $\mathfrak{G}(f)$.

2. Proof of Theorem 1.

2.1. Let M be a compact, oriented manifold and $f \in \mathfrak{F}(M)$. We take $\boldsymbol{R}^2$, the unit circle, and the space of lines through the origin oriented as in 1.3. Let $E = T(M) \mid S_1(f)$ and

$$F = (f^{-1}(T(\mathbf{R}^2))) \mid S_1(f).$$

We define L and G, line bundles over $S_1(f)$ by the exactness of:

$$0 \to L \to E \xrightarrow{df} F \xrightarrow{\pi} G \to 0. \tag{5}$$

The second differential (see [4] for its definition) gives another exact sequence:

$$0 \to T \to E \xrightarrow{d^2f} L^* \otimes G \to 0, \tag{6}$$

where T is the tangent bundle of $S_1(f)$. If we restrict d^2f to L_P we obtain:

$$0 \to R_P \to L_P \xrightarrow{d^2f} (L^* \otimes G)_P \to K_P \to 0, \tag{7}$$

where $\dim R_P = \dim K_P = 1$ (or 0) if $P \in S_1{}^2(f)$ (or $P \in S_1(f) - S_1{}^2(f)$).

Suppose that c is an arc of $S_1(f) - S_1{}^2(f)$. For the rest of this paragraph everything is restricted to c. Let t be an orientation of T; t is a nowhere-zero section in E such that $df(t)$ is a nowhere-zero section of F. Contracting $df(t)$ with the given orientation of F gives another nowhere-zero section h_t of F such that $df(t) \wedge h_t$ is the orientation of F, and $\pi(h_t) = g_t$ is an orientation of G. Since G is a line bundle over c, the orientation g_t gives a trivialization of G. $G \to c \times \mathbf{R}: (P, \lambda g_t(P)) \to (P, \lambda)$. Define Hf: $L \otimes L \to G$ by $Hf(l_1 \otimes l_2) = d^2f(l_1)(l_2)$, the value of $d^2f(l_1) \in \operatorname{Hom}(L, G)$ on l_2. In terms of the trivialization of G, we obtain a mapping:

$$(Hf)_t \colon L \otimes L \to \mathbf{R},$$

which at each point of the arc, c, is a symmetric form of rank $(n-1)$, $(\dim M = n)$ and index τ_t. Clearly $\tau_{-t} = (n - \tau_t - 1)$. Call τ_t the *t-index* of c.

For any point $P \in c$, we can find coordinates $(u, y_1, \cdots, y_{n-1})$ centered at P and coordinates (U, Y) centered at $f(P)$ such that the equations of f are:

$$\begin{cases} U(f(u,y)) = u \\ Y(f(u,y)) = \mathfrak{Q}(y), \end{cases}$$

where $\mathfrak{Q}(y)$ is a non-singular quadratic form of index τ_t, $\dfrac{\partial}{\partial u}$ is a positive multiple of t and $\pi(\dfrac{\partial}{\partial Y} \circ f)$ is a positive multiple of g_t at each point of c in the (u, y)-neighborhood, and $\dfrac{\partial}{\partial U} \wedge \dfrac{\partial}{\partial Y}$ is the orientation of the plane.

2.2. Now suppose $\gamma \in \mathcal{G}(f)$, and c is an arc of $S_1(f) - S_1{}^2(f)$ with orientation t and t-index, τ_t. Let $P \in c$ at which $\gamma \circ f$ is singular. Choose coordinates as in the preceding section, and suppose that in the (U, Y)-neighborhood, γ is given by:

$$\gamma(U, Y) = aY + bU + eU^2 + k(U, Y),$$

where k involves UY, Y^2 and terms of order greater than two. Thus,

$$\gamma \circ f(u, y) = a\mathcal{Q}(y) + bu + eu^2 + k(u, \mathcal{Q}(y)). \tag{8}$$

The order of $k \circ f$ is greater than two in u and y. Since $d(\gamma \circ f)(0, 0) = 0$ we see that $b = 0$. Further since P is a non-degenerate critical point of $\gamma \circ f$, $ae \neq 0$. Since the equation for c in the (u, y)-neighborhood is $y = 0$, we see that P is a relative maximum or minimum of $\gamma \circ f \mid c$. Thus from (8) we can read off the following:

If $\pi(\frac{\partial}{\partial\gamma} \circ f)_P$ is a positive multiple of g_t, then

$$\operatorname{index}_P(\gamma \circ f) = \begin{cases} \tau_t + 1, \text{ if } P \text{ is a relative maximum of } \gamma \circ f \mid c. \\ \tau_t, \text{ if } P \text{ is a relative minimum of } \gamma \circ f \mid c \end{cases}$$

(9)

If $\pi(\frac{\partial}{\partial\gamma} \circ f)_P$ is a negative multiple of g_t, then

$$\operatorname{index}_P(\gamma \circ f) = \begin{cases} n - \tau_t, \text{ if } P \text{ is a relative maximum of } \gamma \circ f \mid c. \\ n - \tau_t - 1, \text{ if } P \text{ is a relative minimum of } \gamma \circ f \mid c \end{cases}$$

2.3. For $\gamma \in \mathcal{G}(f)$, let $\xi(\gamma)$ be the unit tangent at $0 \in \boldsymbol{R}^2$ such that $\xi(\gamma) \wedge \frac{\partial}{\partial\gamma}(0)$ is the orientation of $\boldsymbol{R}^2$ and $\xi(\gamma)$ is tangent to the level line of γ through the origin.

Given an arc c of $S_1(f) - S_1{}^2(f)$, oriented by t, define $\theta_t \colon c \to S^1 \subset \boldsymbol{R}^2$ by:

$$\theta_t(P) = \pi_2\left(\frac{f_* t(P)}{\| f_* t(P) \|}\right),$$

where $\pi_2 \colon T(\boldsymbol{R}^2) = \boldsymbol{R}^2 \times \boldsymbol{R}^2 \to \boldsymbol{R}^2$, projection onto the fibre at the origin. The critical set of $\gamma \circ f$ on c is just $\theta_t^{-1}(\pm \xi(\gamma))$. We relate the number of times that the points $\pm(\xi(\gamma))$ are covered by θ_t to the number of critical points of $\gamma \circ f$ on c of various indices.

Let $N_c(\sigma)$ = the number of points on c at which $\gamma \circ f$ is singular and has index σ. Suppose c is oriented and S^1 is an oriented circle, let $\beta \colon c \to S^1$.

For any regular value, x, of β we define $n(x;\beta)$ to be the number of β-preimages of x. Here $P \in \beta^{-1}(x)$ is counted positively (negatively) if β is orientation preserving (reversing) at P.

LEMMA. *Let c be a component of $S_1(f) - S_1{}^1(f)$ with orientation t, and let $\gamma \in \mathfrak{G}(f)$. Then $\pm\xi(\gamma)$ are regular values of θ_t and*

$$n(\xi(\gamma);\theta_t) = N_c(\tau_t) - N_c(\tau_t + 1)$$

and

$$n(-\xi(\gamma);\theta_t) = N_c(n - \tau_t) - N_c(n - \tau_t - 1).$$

Proof. Here it is of course understood that $n(\cdot\,;\theta_t)$ is computed taking t as the orientation of c and S^1 with the usual orientation (see 1.3) of the unit circle in $\boldsymbol{R}^2$.

It is clear that $\pm\xi(\gamma)$ are regular values of θ_t if and only if zero is a regular value of $\gamma \circ \pi_2 \circ f_* \circ t\colon c \to \boldsymbol{R}$. But the second statement is equivalent to $\gamma \circ f \mid c$ has only non-degenerate singularities which is true since $\gamma \in \mathfrak{G}(f)$. Further it is readily checked that if $\theta_t(P) = \xi(\gamma)$ then θ_t is orientation preserving at P if and only if $\gamma \circ f \mid c$ has a relative minimum at P, and if $\theta_t(P) = -\xi(\gamma)$, θ_t is orientation preserving at P if and only if $\gamma \circ f \mid c$ has a relative maximum at P.

Also note that $\theta_t(P)$ is a positive (negative) multiple of $\xi(\gamma)$ if and only if $\pi(\frac{\partial}{\partial\gamma} \circ f)(P)$ is a positive (negative) multiple of $g_t(P)$. In fact suppose $\theta_t(P) = \epsilon\xi(\gamma)$ and $\pi(\frac{\partial}{\partial\gamma} \circ f)(P) = \frac{1}{\mu}\, g_t(P)$. The second statement is equivalent to $df(t(P)) \wedge \mu(\frac{\partial}{\partial\gamma} \circ f)(P)$ orients F_P. This in turn is equivalent to $\mu\pi_2 f_* t(P) \wedge \frac{\partial}{\partial\gamma}(0)$ orients $\boldsymbol{R}^2$. But

$$\mu\pi_2 f_* t(P) \wedge \frac{\partial}{\partial\gamma}(0) = \mu\epsilon \, \| f_* t(P) \| \, \xi(\gamma) \wedge \frac{\partial}{\partial\gamma}(0)$$

and $\xi(\gamma) \wedge \frac{\partial}{\partial\gamma}(0)$ also orients $\boldsymbol{R}^2$, so $\mu\epsilon > 0$.

Thus we can make the following translation of (9) which completes the proof of the lemma:

If $\theta_t(P) = \xi(\gamma)$, then

$$\operatorname{index}_P(\gamma \circ f) = \begin{cases} \tau_t, \text{ if } \theta_t \text{ is orientation preserving at } P. \\ \tau_t + 1, \text{ if } \theta_t \text{ is orientation reversing at } P. \end{cases}$$

(10)

If $\theta_t(P) = -\xi(\gamma)$, then

$$\operatorname{index}_P(\gamma \circ f) = \begin{cases} n - \tau_t, \text{ if } \theta_t \text{ is orientation preserving at } P. \\ n - \tau_t - 1, \text{ if } \theta_t \text{ is orientation reversing at } P. \end{cases}$$

Let $\alpha\colon S^1 \to P(\boldsymbol{R}^2) : x \to [x]$ be the usual double covering of the space of lines through the origin of $\boldsymbol{R}^2$ by the circle of directions at the origin of $\boldsymbol{R}^2$. The map $\alpha \circ \theta_t$ is just $\phi \mid c$, where ϕ was defined on all of $S_1(f)$ in 1.3.

COROLLARY. $n([\xi(\gamma)]; \phi \mid c)$

$$= N_c(\tau_t) + N_c(n - \tau_t) - N_c(\tau_t + 1) - N_c(n - \tau_t - 1).$$

2.4. From now on until the end of the proof of the theorem, we assume that $n = \dim M$ is even.

LEMMA. *Let $f \in \mathfrak{F}(M)$, then there is a unique orientation t of $S_1(f)$ such that if c is any component of $S_1(f) - S_1{}^2(f)$, the t-index of c is even.*

The theorem follows from this lemma very easily. Take $\gamma \in \mathfrak{G}(f)$ so that the singular set of $\gamma \circ f$ is in $S_1(f) - S_1{}^2(f)$. The subset of $\mathfrak{G}(f)$ with this property is just the complement of a finite number of lines in $\mathfrak{G}(f)$. That is, we delete from $\mathfrak{G}(f)$ any element whose level lines are parallel to $\phi(P)$ for some $P \in S_1{}^2(f)$. This restriction of the choice of γ is inessential; it just facilitates counting. Let $N(\tau) =$ (the number of points at which $\gamma \circ f$ has index τ) $= \sum_c N_c(\tau)$, where the summation runs over the components of $S_1(f) - S_1{}^2(f)$. By (2),

$$\chi(M) = \sum_c \left(\sum_\tau (-1)^\tau N_c(\tau) \right).$$

If c is any component of $S_1(f) - S_1{}^2(f)$, the corollary of 2.3 yields:

$$\sum_\tau (-1)^\tau N_c(\tau) = n([\xi(\gamma)]; \phi \mid c).$$

If C is any component of $S_1(f)$, $r(C)$, the degree of $\phi \mid C$ is just $\sum_c n(\xi(\gamma); \phi \mid c)$, where the sum runs over the components of $C - S_1{}^2(f)$. Thus we obtain our result:

$$\chi(M) = \sum_C r(C).$$

2.5. We now prove the lemma stated in 2.4. To give an orientation to $S_1(f)$, we must give one to each component. If a component has no cusps, give it the orientation t so that t-index of the component is even. This can always be done and in only one way since n is even.

Now suppose C is a component of $S_1(f)$ with $S_1{}^2(f) \cap C \neq \emptyset$. In Sec. 3.4 of [4], it is shown that there is a unique orientation t of C such that for any $P \in S_1{}^2(f) \cap C$ if f is put in normal form (4) at P with the additional requirements on the coordinates that $\frac{\partial}{\partial x}$ at P is a positive multiple of $t(P)$ and that $\frac{\partial}{\partial U} \wedge \frac{\partial}{\partial Y}$ agrees with the given orientation of $\boldsymbol{R}^2$, then the coefficient a of xu is $(-1)^{\tau+1}$ where τ is the index of the quadratic form in the z's. It is this orientation of C which we want.

Let $P \in S_1{}^2(f) \cap C$, and suppose t is the orientation of C described above. We show that the two arcs which abut P have even t-index. The normal form of f at P is:

$$\left\{\begin{array}{l} U \circ f(u,x,z) = u \\ Y \circ f(u,x,z) = \mathfrak{Q}(z) + axu + bx^3 \end{array}\right. \tag{11}$$

where $\mathfrak{Q}(z)$ is a non-singular quadratic form of index τ, $\frac{\partial}{\partial x}$ at P is a positive multiple of $t(P)$, $\frac{\partial}{\partial U} \wedge \frac{\partial}{\partial Y}$ agrees with the given orientation of $\boldsymbol{R}^2$ and $a = (-1)^{\tau+1}$. In these coordinates, $S_1(f)$ is given by $\{z_1 = \cdots = z_{n-2} = 0,\ au + 3bx^2 = 0\}$. We will work entirely in these coordinate neighborhoods. Let $Q \in S_1(f) \cap C - S_1{}^2(f)$, then the (x,u,z) coordinates of Q are $(s, \frac{-3b}{a}s^2, 0)$. We change coordinates so that Q and $f(Q)$ are centers. In particular let:

$$\left\{\begin{array}{l} \bar{u} = u + \dfrac{3b}{a}s^2 \\ \bar{x} = x - s \\ \bar{z} = z \end{array}\right. \quad \text{and} \quad \left\{\begin{array}{l} \bar{U} = U + \dfrac{3b}{a}s^2 \\ \\ \bar{Y} = Y - bs^3 - asU. \end{array}\right.$$

Substituting in (11) and dropping the bars, we obtain:

$$\left\{\begin{array}{l} U \circ f(u,x,z) = u \\ Y \circ f(u,x,z) = \mathfrak{Q}(z) + 3bsx^2 + axu + bx^3. \end{array}\right.$$

The equations of $S_1(f)$ are $\{z = 0,\ au + 3bx(x + 2s) = 0\}$. The vector field $(\frac{\partial}{\partial x} - \frac{6b}{a}(s + x)\frac{\partial}{\partial u})$ is tangent to $S_1(f)$ and is a positive multiple of t on C. The value of this vector field at Q has as f_* image: $\frac{-6bs}{a}\frac{\partial}{\partial U}$. Thus

$g_t(Q)$ is a positive multiple of $\pi(\frac{-6bs}{a}\frac{\partial}{\partial Y})$. So the index of the arc of C containing Q is the same as the index of $\frac{-6bs}{a}(\mathfrak{Q}(z)+3bsx^2)$, which is the same as the index of $(-1)^{\tau}(bs\mathfrak{Q}(z)+x^2)$, where τ is the index of $\mathfrak{Q}(z)$. Thus if τ is even, the arc of C containing Q has index τ if $bs>0$ and $n-2-\tau$ if $bs<0$. If τ is odd, the arc of C containing Q has index $n-2-\tau+1$ if $bs>0$, and $\tau+1$ if $bs<0$. In all cases the arcs abutting the cusp have even indices relative to the chosen orientation of C.

BRANDEIS UNIVERSITY,
WALTHAM, MASSACHUSETTS.

REFERENCES.

[1] N. Kuiper, "Immersions with minimal total absolute curvature," *Centre Belge des Recherches Mathematiques* (1958), pp. 75-88.

[2] H. Levine, *Singularities of differentiable mappings*, duplicated notes, Mathematics Institute, Bonn University (1959).

[3] ———, "The singularities, S_1^q," *Illinois Journal of Mathematics*, vol. 8 (1964), pp. 152-168.

[4] ———, "Elimination of cusps," *Topology*, vol. 3 (1964), pp. 263-296.

[5] E. Pitcher, "Inequalities of critical point theory," *Bulletin of the American Mathematical Society*, vol. 64 (1958), pp. 1-30.

[6] R. Thom, "Un lemma sur les applications differentiables," *Boletin de la Société Matematicae de Mexico*, Segunda Serie, vol. 1 (1956), pp. 59-71.

[7] ———, "Les singularites des applications differentiables," *Annales de l'Institut Fourier*, vol. 4 (1955-1956), pp. 43-87.

[8] H. Whitney, "On the singularities of mappings of Euclidean spaces I. Mappings of the plane into the plane," *Annals of Mathematics*, vol. 62 (1955), pp. 374-410.

THE EXTENSION PROBLEM IN COMPLEX ANALYSIS II; EMBEDDINGS WITH POSITIVE NORMAL BUNDLE.

By Phillip A. Griffiths.

The extension problem we have in mind is simply to extend certain "analytic objects" from a compact analytic submanifold X of a complex manifold W to a neighborhood of X in W. The analytic objects are to be such things as holomorphic vector bundles, cohomology classes, or holomorphic mappings. This problem arises naturally when one considers the following question: How much about W can we determine knowing X and the normal bundle of X in W? In [7] we have discussed the case of a trivial normal bundle; here we shall mainly be concerned with the situation when the normal bundle is positive. The main results of this paper are stated in § II.

Now the extension problem falls naturally into two pieces. On is the formal problem which is easy to treat using sheaf theory. The point is that the obstructions to extending analytic objects are cohomology classes with coefficients in a coherent analytic sheaf which involves the symmetric powers of the dual normal bundle (§ 1). The other part of the extension problem is the question of existence. Here our approach has been to use geometric arguments (based on projective embeddings and the theory of continuous systems) to reduce the existence question to linear problems (§ III). By use of monoidal transformations (§ IV) we are able to reduce the linear problems to proving a "vanishing theorem," an open manifolds, for cohomology with coefficients in the sheaf of germs of holomorphic cross sections of an analytic vector bundle. Our technique for this is to use the theory of harmonic integrals developed by Joe Kohn [15] together with a systematic treatment of Hermitian differential geometry (§§ VI and VII). In § VIII we complete the proof of our vanishing theorem.

The central part of the present paper is, we feel, the various differential-geometric calculations which give a precise analytic reflection of the geometric assumptions in the problem. As one example, in the appendix to § IV we use curvature arguments to characterize positive vector bundles as those for which a suitable set of integrals over subvarieties are positive. Again, in § V,

---

Received April 14, 1965.

we use curvature calculations to prove the following result: If $X \subset W$ is a non-singular hypersurface in a closed algebraic variety W, if $X \subset W$ looks as though it should be the monoidal transform of some $X' \subset W'$, and if the (hypothetical) normal bundle of $X' \subset W'$ is positive, then X' and W' exist and $X \subset W$ is the monoidal transformation along X' in W'. Finally, curvature arguments lead to two integral formulae (§§ VII. 3 and VII. 5) which, among other things, yield the required vanishing theorems.

To close the introduction, we should like to remark on the question which was the original motivation for the present study. This is the rigidity problem, of Nirenberg and Spencer [17], which goes as follows: Given a germ of embedding $X \subset W (\dim X \geqq 2)$ with positive normal bundle, is W determined by a finite neighborhood of X? Hironaka has answered this question affimatively for hypersurfaces, and we give also an affirmative answer in general, provided that the normal bundle of X in W is sufficiently positive (geometrically, the complete continuous system generated by X must be large).

Concerning notation, $X, W, Y, \cdots$ are non-singular complex manifolds and $\mathfrak{O}_X, \mathfrak{O}_W, \mathfrak{O}_Y, \cdots$ are the structure sheaves. (As usual, we write $\mathfrak{O}$ for $\mathfrak{O}_X$ if no confusion is possible.) A holomorphic vector bundle over X will be denoted by $\boldsymbol{E} \to X$; $\Omega(\boldsymbol{E})$ is the sheaf of germs of holomorphic sections of $\boldsymbol{E}$. We let $\boldsymbol{T}(X)$ be the holomorphic tangent bundle of X and $A^{p,q}(\boldsymbol{E})$ the sheaf of germs of C^∞ sections of $\boldsymbol{E} \otimes \wedge^p \boldsymbol{T}(X)^* \otimes \wedge^q \overline{\boldsymbol{T}(X)}^*$ (here $\boldsymbol{E}^*$ is the dual bundle to $\boldsymbol{E}$). Also we set $A^q(\boldsymbol{E}) = A^{0,q}(\boldsymbol{E})$. The other notations are the standard ones from [9].

Outline

I. The Extension Problem for Analytic Objects.
1. Germs of Embeddings and the System of Jet Neighborhoods of a Submanifold
2. Extension Problem for Analytic Objects
3. Obstructions to the Formal Extension Problem
4. The Obstructions to Uniqueness
5. Some Special Cases
    (a) Line Bundles
    (b) Transverse Foliation
    (c) Rigidity of Holomorphic Embeddings

II. The Normal Bundle and the Extension Problem.
1. Some Implications of the Extension and Uniqueness Problems for Cohomology

2. Positive and Negative Holomorphic Bundles
3. Statement of the Main Results
4. Outline of the Proof of Theorem III.

III. Positive Embeddings and the Proof of Theorem II.

1. A Simple Example of Rigidity
2. On the Theory of Continuous Systems
3. Positive Embeddings and Continuous Systems
4. Proof of Theorem II from Theorem I

IV. Some Examples and Remarks about Positive Embeddings.

1. Proof of Theorem I for Positively Embedded Hypersurfaces
2. Rigidity of Some Special Embeddings
   (i) Projective Embeddings
   (ii) Embeddings in Abelian Varieties
3. On Ample Hypersurfaces
4. A Partial Generalization to Higher Codimension

Appendix to § IV: Numerically Positive Bundles.

A. 1. The Cone of Positive Polynomials for Ample Bundles
A. 2. Numerically Positive Bundles
A. 3. Proof of Theorem IV

V. Positive Embeddings and Monodial Transformations.

1. Monoidal Transformations
2. Construction of a Blowing Down Map
3. Reduction to the Codimension One Case

VI. Differential-Geometric Foundations.

1. Local Hermitian Geometry
2. Holomorphic Vector Bundles
3. Induced Connexions in Sub-and-Quotient Bundles
4. Comparison of Hermitian Forms
5. Positive and Negative Bundles

VII. Integral Formulae and Vanishing Theorems for Complex Manifolds.

1. Harmonic Integrals on Open Manifolds
2. Statement of the Vanishing Theorems
3. A Special Integral Formula
4. Some Remarks and Applications
   (i) Compact Manifolds
   (ii) Open Manifolds

5. An Integral Formula for Functions on Complex Manifolds
6. Proof of Theorem VI

VIII. Properties of Tubular Neighborhoods and the Proof of Theorem III.
1. Metrics in Line Bundles
2. Continuous Systems and the Pseudo-Characteristics of Tubular Neighborhoods
3. Another Proof of the Concavity-Convexity of Tubular Neighborhoods
4. Proof of Theorem VII.

Bibliography.

I. The Extension Problem for Analytic Objects.

1. Germs of embeddings and the system of jet neighborhoods of a Submanifold. Let X be a compact, complex manifold and consider an embedding $\phi: X \to W$ of X in a complex manifold W. We shall consider another such embedding $\phi': X \to W'$ as equivalent to $\phi: X \to W$ if there exist open neighborhoods U of $\phi(X)$ in W and U' of $\phi'(X)$ in W' and a bi-holomorphic mapping $f: U \to U'$ such that

$$\begin{tikzcd} & U \arrow[d, "f"] \\ X \arrow[ur, "\phi"] \arrow[r, "\phi'"'] & U' \end{tikzcd}$$

commutes. Thus we are considering *germs of embeddings,* and we shall usually suppress the mapping ϕ and write $X \subset W$ as a representative of the germ of embedding defined by $\phi: X \to W$. If we let $\mathfrak{O}_W$ be the structure sheaf of W and set $\mathfrak{O}_* = \mathfrak{O}_W \mid X$, then the pair $(X, \mathfrak{O}_*)$ forms a generalized complex space X_*. Clearly X_* depends only on, and in fact determines, the germ of embedding defined by $X \subset W$.

Let $\mathfrak{J} \subset \mathfrak{O}_W$ be the ideal sheaf of X. Then $\mathfrak{O}_W/\mathfrak{J} = \mathfrak{O}$ is a sheaf supported on X and is, in fact, the structure sheaf of X. For each $\mu \geqq 0$, we let $\mathfrak{O}^{(\mu)} = \mathfrak{O}_W/\mathfrak{J}^{\mu+1}$. The pair $(X, \mathfrak{O}^{(\mu)}) = X_\mu$ forms again a generalized complex space which may be thought of as the neighborhood of order μ, or *μ-neighborhood,* of $X \subset W$. By the *stable neighborhood* of $X \subset W$, we mean the μ-neighborhood for large μ.

There are obvious inclusions $X_\mu \to X_{\mu+\nu}$ $(\nu \geqq 0)$ which lead to maps $\mathfrak{O}^{(\mu+\nu)} \to \mathfrak{O}^{(\mu)}$. The inverse limit sheaf $\hat{\mathfrak{O}} = \underset{\leftarrow \mu}{\lim} \mathfrak{O}^{(\mu)}$ then exists, and the pair $(X, \hat{\mathfrak{O}})$ forms again a complex space called the *formal neighborhood* determined by $X \subset W$.

2. Extension problem for analytic objects. Given $X \subset W$, we have in mind extending certain "analytic objects" from X to W. By an *analytic object*, we mean something along the following lines: (i) an analytic vector bundle $\boldsymbol{E} \to X$, (ii) an element $\sigma \in H^q(X, \Omega(\boldsymbol{F}))$ where $\boldsymbol{F} \to X$ is the restriction of a holomorphic vector bundle $\boldsymbol{F}^\# \to W$; or (iii) a holomorphic mapping $f \colon X \to Z$ for some complex manifold Z. To *extend an analytic object* α on X means to find α_* on X_* such that $\alpha_* \mid X = \alpha$; to *formally extend* α means that there exists $\hat{\alpha}$ on $\hat{X}$ with $\hat{\alpha} \mid X = \alpha$. Finally, the *stable extension problem* is the following: Given α on X, can we find a $\mu_0 = \mu_0(\alpha)$ such that, for any germ of embedding $X \subset W'$ for which $X'_\mu \cong X_\mu$ and where there exists α'_μ on X'_μ with $\alpha'_\mu \mid X = \alpha$ $(\mu \geqq \mu_0)$, there exists α'_* on X'_* with $\alpha'_* \mid X = \alpha$? Similarly, we may speak of the *formal stable extension problem.* Throughout, the adjective "stable" should be thought of as meaning "finitely many obstructions."

Now whenever the extension problem is defined, we may speak of the *uniqueness problem* (if α_* and β_* are extensions of α, then does $\alpha_* = \beta_*$?) and the *formal uniqueness problem.* More generally, we may speak of the *stable uniqueness problem* (if $\alpha_* \mid X_\mu = \beta_* \mid X_\mu$, then does $\alpha_* = \beta_*$ provided that $\mu \geqq \mu_0$?) and the *formal stable uniqueness problem.*

After this list of definitions, it is perhaps well to state the general problem which has served as a guideline: Given an algebraic variety W and a subvariety $X \subset W$, how much can we say about W on the basis of the information given by X_μ for finite μ?

3. The obstructions to the formal extension problem. Given $X \subset W$, there is a sequence of inclusions $X \subset X_1 \subset \cdots \subset X_\mu \subset X_{\mu+1} \subset \cdots$. Given then an analytic object α on X, we may try to extend it step by step; if this can be done, then we have solved the formal extension problem for α. If we have found an extension $\alpha_{\mu-1}$ of α to $X_{\mu-1}$, then there will be a cohomology class $\omega(\alpha_{\mu-1})$ which will be the μ*-th* obstruction to extending α. If all $\omega(\alpha_{\mu-1}) = 0$, then we may formally extend α.

Before stating explicitly where $\omega(\alpha_{\mu-1})$ belongs, we introduce the following notations: $\boldsymbol{N} \to X$ is the *normal bundle* of $X \subset W$, $\boldsymbol{N}^*$ is the dual bundle, and $\boldsymbol{N}^{*\mu}$ is the μ-th symmetric product of $\boldsymbol{N}^*$. For any holomorphic vector bundle $\boldsymbol{F} \to X$, $\boldsymbol{F}(\mu) = \boldsymbol{F} \otimes \boldsymbol{N}^{*\mu}$. Finally, $\boldsymbol{T} = \boldsymbol{T}(W) \mid X$, $\boldsymbol{T}(W)$ being the tangent bundle of W. We have the usual exact bundle sequence

$$0 \to \boldsymbol{T}(X) \to \boldsymbol{T} \to \boldsymbol{N} \to 0. \tag{1.1}$$

PROPOSITION 1.1. *If* α *is a holomorphic vector bundle* $\boldsymbol{E}\rightarrow X$, *then* $\omega(\alpha_{\mu-1})\in H^2(X,\Omega(\mathrm{Hom}(\boldsymbol{E},\boldsymbol{E})(\mu)))$.

PROPOSITION 1.2. *If* α *is an element in* $H^q(X,\Omega(\boldsymbol{F}))$, *then* $\omega(\alpha_{\mu-1})\in H^{q+1}(X,\Omega(\boldsymbol{F}(\mu)))$.

PROPOSITION 1.3. *If* α *is a holomorphic mapping* $f\colon X\rightarrow Z$, *then* $\omega(\alpha_{\mu-1})\in H^1(X,\boldsymbol{T}^{\#}(\mu)))$ *where* $\boldsymbol{T}^{\#}=f^{-1}(\boldsymbol{T}(Z))$.

Proof of Proposition 1.2. Over X we have

$$0\rightarrow\Omega(N^{*\mu})\rightarrow\mathfrak{O}^{(\mu)}\rightarrow\mathfrak{O}^{(\mu-1)}\rightarrow 0 \tag{1.2}$$

which immediately gives

$$0\rightarrow\Omega(\boldsymbol{F}(\mu))\rightarrow\mathfrak{O}^{(\mu)}(\boldsymbol{F}^{\#})\rightarrow\mathfrak{O}^{(\mu-1)}(\boldsymbol{F}^{\#})\rightarrow 0 \tag{1.3}$$

where $\mathfrak{O}^{(\mu)}(\boldsymbol{F}^{\#})=\mathfrak{O}^{(\mu)}\otimes_{\mathfrak{O}_*}\Omega(\boldsymbol{F}^{\#})$ are the sections up to order μ of $\boldsymbol{F}^{\#}$ along X. The cohomology sequence of (1.3) gives

$$H^q(\mathfrak{O}^{(\mu)}(\boldsymbol{F}^{\#}))\rightarrow H^q(\mathfrak{O}^{(\mu-1)}(\boldsymbol{F}^{\#}))\xrightarrow{\delta}H^{q+1}(\Omega(\boldsymbol{F}(\mu)));$$

clearly, $\alpha_{\mu-1}\in H^q(\mathfrak{O}^{(\mu-1)}(\boldsymbol{F}^{\#}))$ and the obstruction to extending $\alpha_{\mu-1}$ is $\delta(\alpha_{\mu-1})\in H^{q+1}(\Omega(\boldsymbol{F}(\mu)))$.

Proof of Proposition 1.3. We first claim that it will suffice to assume that $f\colon X\rightarrow Z$ is an embedding. Indeed, replace Z by $Z'=W\times Z$ and f by f' where $f'(x)=(x,f(x))$. If we can find an extension F' of f', then $F'(w)=(g(w),F(w))$ and F will be an extension of f. Furthermore, $f'^{-1}(\boldsymbol{T}(Z'))=\boldsymbol{T}\oplus\boldsymbol{T}^{\#}$ and so, if Proposition 1.3 holds for embeddings, then the obstructions to extending F' will lie in $H^1(\Omega(\boldsymbol{T}(\mu)))\oplus H^1(\Omega(\boldsymbol{T}^{\#}(\mu)))$. Finally, if we require that the first factor g of F' should be the identity, $g(w)=w$, then the obstructions to finding F will clearly lie in $H^1(\Omega(\boldsymbol{T}^{\#}(\mu)))$. Thus we may suppose that f is an embedding and we may even require that the extension F of f should be an embedding of W.

Suppose then that we have $f_{\mu-1}\colon X_{\mu-1}\rightarrow Z$ and assume, for simplicity,* that $\mu\geqq 2$. We first choose a coordinate covering $\{U_\alpha\}$ of W by polycylinders with coordinates $(z_\alpha,w_\alpha)=(z_\alpha^1,\cdots,z_\alpha^n;w_\alpha^1,\cdots,w_\alpha^r)$ in U_α such that $U_\alpha\cap X=\{w_\alpha=0\}$. Then $f_{\mu-1}\mid U_\alpha=f_\alpha$ gives an embedding of U_α into Z, and so we may cover Z by polycylinders V_α with coordinates $(z_\alpha,w_\alpha,\xi_\alpha)$ in V_α such that $f_\alpha(z_\alpha,w_\alpha)=(z_\alpha,w_\alpha,0)$. Now

---

* The case $\mu=1$ is given in 1.5, section (C) below.

$$(f_\alpha \circ f_\beta^{-1})(z_\beta, w_\beta) = (z_\alpha(z_\beta, w_\beta), w_\alpha(z_\beta, w_\beta), \xi_\alpha(z_\beta, w_\beta))$$

is a power series of the form

$$\sum_{\mu_1+\cdots+\mu_r=\mu} \Lambda_{\alpha\beta\mu_1\cdots\mu_r}(z_\beta) w_\beta^{\mu_1} \cdots w_\beta^{\mu_r} + [\mu+1]$$

where the $\Lambda_{\alpha\beta\mu\cdots\mu_r}(z_\beta)$ are vector valued functions and $[\mu+1]$ vanishes to order $\mu+1$ in w_β. Then $\Lambda = \{\Lambda_{\alpha\beta_{\mu_1\cdots\mu_r}}\}$ gives a 1-cochain with values in

$$\Omega(\boldsymbol{T}(Z) \mid f(X)) \otimes \mathfrak{J}^\mu/\mathfrak{J}^{\mu+1} \cong \Omega(\boldsymbol{T}^\#(\mu)).$$

It is, of course, straightforward to check that $\delta\Lambda = 0$; that $\Lambda - \Lambda' = \delta\lambda$ if, by a different coordinatization, we arrive at another 1-cochain Λ'; and that, if $\Lambda = \delta\Phi$, then we may lift $f_{\mu-1}$ to $f_\mu : X_\mu \to Z$. Clearly then $\omega(f_{\mu-1}) = \Lambda \in H^1(\Omega(\boldsymbol{T}^\#(\mu)))$.

We shall omit the proof of Proposition 1.1.

4. The obstructions to uniqueness. From § 1.3 above, it is fairly clear that the obstructions to uniqueness are cohomology classes in dimension one less than the obstructions to extension. However, uniqueness must be interpreted in an individual manner for each type of analytic object. For cohomology classes, the meaning of uniqueness of extension is obvious; for a holomorphic mapping $f: X \to Z$, uniqueness means that the power series expansion of an extension $F: W \to Z$ is uniquely determined. For bundles, to say that two extensions $\boldsymbol{E}^\# \to W$ and $\boldsymbol{E}^{\#\prime} \to W$ of $\boldsymbol{E} \to W$ are the same means that there is a bundle equivalence

$$\begin{tikzcd} G: \boldsymbol{E}^\# \arrow[r] \arrow[d] & \boldsymbol{E}^{\#\prime} \arrow[d] \\ W \arrow[r] & W. \end{tikzcd}$$

PROPOSITION 1.4. *Referring to Propositions* 1.1-1.3, *the obstructions to uniqueness for the extension problem are classes in* $H^1(\Omega(\operatorname{Hom}(\boldsymbol{E},\boldsymbol{E})(\mu)))$, $H^q(\Omega(\boldsymbol{F}(\mu)))$, *and* $H^0(\Omega(\boldsymbol{T}^\#(\mu)))$ *respectively.*

5. Same special cases.

(a) Line bundles. We use the notation $\Omega(\mu) = \Omega(\boldsymbol{N}^{*\mu})$. The μ-th obstruction to extending a line bundle $\boldsymbol{E} \to X$ is an element in $H^2(X, \Omega(\mu))$. In order to check this directly, we let $\mathfrak{O}_X$, $\mathfrak{O}_W$ be the structure sheaves of X, W and $\mathfrak{O}_X{}^*$, $\mathfrak{O}_W{}^*$ the corresponding sheaves of units. Denote by $\mathfrak{J}^* \subset \mathfrak{O}_W{}^*$ the germs of the form $\exp(2\pi\sqrt{-1}\,f)$ where $f \in \mathfrak{J}$. Then, considering $X \subset W$ as a germ of embedding, $\exp: \mathfrak{J} \to \mathfrak{J}^*$ is an isomorphism, and we get a commutative diagram

$$\begin{tikzcd}
& & 0 \arrow[d] & 0 \arrow[d] & \\
& 0 \arrow[r] \arrow[d] & \mathfrak{J} \arrow[r] \arrow[d] & \mathfrak{J}^* \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathbf{Z} \arrow[r] \arrow[d] & \mathfrak{O}_W \mid X \arrow[r] \arrow[d] & \mathfrak{O}_W^* \mid X \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathbf{Z} \arrow[r] \arrow[d] & \mathfrak{O}_X \arrow[r] \arrow[d] & \mathfrak{O}_X^* \arrow[r] \arrow[d] & 0 \\
& 0 & 0 & 0 &
\end{tikzcd}$$

The extension problem for line bundles is to lift $\xi \in H^1(\mathfrak{O}_X^*)$ to $H^1(\mathfrak{O}_W^* \mid X)$. From the exact cohomology diagram of (1.4), the obstruction to lifting ξ is in $H^2(\mathfrak{J}^*)$. Now $\mathfrak{J}$ has a *decreasing filtration* by the subsheaves $\mathfrak{J}^\mu$ $(\mu \geqq 1)$, and the *associated graded sheaf* is $\sum_{\mu \geqq 1} \Omega(\mu)$. Thus, the μ-th obstruction to extending ξ is in $H^2(\Omega(\mu))$.

Slightly more generally, we have

PROPOSITION 1.5. *Let Z be a complex manifold and let $Y \subset Z$ be either a complex submanifold or an open subset. Assume that we have the sequences*

$$\left\{\begin{array}{l} 0 \to H^2(Z, \mathfrak{O}_Z) \to H^2(Y, \mathfrak{O}_Y) \\ \quad H^2(Z, \mathbf{Z}) \ \to H^2(Y, \mathbf{Z}) \to 0 \\ \quad H^1(Z, \mathfrak{O}_Z) \to H^1(Y, \mathfrak{O}_Y) \to 0. \end{array}\right. \tag{1.5}$$

Then every line bundle on Y extends to Z.

Proof. This follows from the exact diagram

$$\begin{tikzcd}
H^1(Z, \mathfrak{O}_Z) \arrow[r] \arrow[d] & H^1(Z, \mathfrak{O}_Z^*) \arrow[r] \arrow[d] & H^2(Z, \mathbf{Z}) \arrow[r] \arrow[d] & H^2(Z, \mathfrak{O}_Z) \arrow[d] \\
H^1(Y, \mathfrak{O}_Y) \arrow[r] & H^1(Y, \mathfrak{O}_Y^*) \arrow[r] & H^2(Y, \mathbf{Z}) \arrow[r] & H^2(Y, \mathfrak{O}_Y).
\end{tikzcd} \tag{1.6}$$

(b) Transverse foliation. An embedding $X \subset W$ is called *transversely foliated* if, through each $x \in X$, there passes a complex subvariety, of dimension complementary to X, $S_x \subset W$ such that $S_x \cap X = \{x\}$ and $S_x \cap S_{x'} = \phi$ for $x \neq x'$.

PROPOSITION 1.6. *The obstructions to transversely foliating $X \subset W$ are classes in $H^1(X, \Omega(\mathbf{T}(X)(\mu)))$.*

Proof. $X \subset W$ is transversely foliated if, and only if, there exists a holomorphic map $F\colon W \to X$ such that $F \mid X =$ identity. The obstructions to finding such an F are in $H^1(\Omega(\mathbf{T}(X)(\mu)))$ since, in this case, $\mathbf{T}^\# = \mathbf{T}(X)$.

Remark. The first obstruction to transversely foliating $X \subset W$ is the obstruction to holomorphically splitting the exact bundle sequence (1.1).

This latter obstruction is well-known to be a class in $H^1(X, \Omega(\mathrm{Hom}(\boldsymbol{N}, \boldsymbol{T}(X)))$. This checks Proposition 1.6 at the first level.

(c) **Rigidity of holomorphic embeddings.** The *rigidity question* for germs of holomorphic embeddings is simply the following: Is the germ of embedding defined by $\phi\colon X \rightarrow W$ determined by a finite neighborhood X_μ? More precisely, does there exist a μ_0 such that, given $\phi\colon X \rightarrow W$, $\phi'\colon X \rightarrow W'$, and an equivalence $f_\mu\colon X_\mu \rightarrow X'_\mu$ $(\mu \geqq \mu_0)$, there exists an equivalence of germs $f_*\colon X_* \rightarrow X'_*$? From Proposition 1.3 we get

PROPOSITION 1.7. *If there exists a μ_0 such that $H^1(X, \Omega(\boldsymbol{T}(\mu))) = 0$ for $\mu \geqq \mu_0$, then the germ of embedding defined by $X \subset W$ is formally rigid. If furthermore $H^0(X, \Omega(\boldsymbol{T}(\mu))) = 0$ for $\mu \geqq \mu_0$, then any formal equivalence $\hat{f}\colon \hat{X} \rightarrow \hat{X}'$ is uniquely determined by $f_{\mu_0} = \hat{f} \mid X_{\mu_0}$.*

Remark. Suppose that we have embeddings $\phi\colon X \rightarrow W$, $\phi'\colon X \rightarrow W'$ with normal bundles $\boldsymbol{N}, \boldsymbol{N}'$ respectively. Assume that we have a bundle isomorphism

$$\begin{tikzcd}
\Psi\colon \boldsymbol{N} \arrow[r] \arrow[d] & \boldsymbol{N}' \arrow[d] \\
\Psi\colon X \arrow[r] & X.
\end{tikzcd}$$

We want to know when Ψ is induced by an equivalence on 1-neighborhoods $\Lambda\colon X_1 \rightarrow X'_1$. Thus we wish to fill in the middle column in the diagram

(1.7)

$$\begin{tikzcd}
0 \arrow[r] & \boldsymbol{T}(X) \arrow[r, "\alpha"] & \boldsymbol{T}' \arrow[r, "\beta"] & \boldsymbol{N}' \arrow[r] & 0 \\
0 \arrow[r] & \boldsymbol{T}(X) \arrow[r, "\gamma"] \arrow[u, "\Psi_*"'] & \boldsymbol{T} \arrow[r] \arrow[u, "\Lambda_1"] \arrow[ur, phantom, "\sigma"] & \boldsymbol{N} \arrow[r] \arrow[u, "\Psi"'] & 0.
\end{tikzcd}$$

From (1.7) we get a diagram

(1.8)

$$\begin{tikzcd}
 & & 0 \arrow[d] & 0 \arrow[d] \\
 & & H^0(\Omega(\mathrm{Hom}(\boldsymbol{N}, \boldsymbol{T}'))) \arrow[r] \arrow[d] & H^0(\Omega(\mathrm{Hom}(\boldsymbol{N}, \boldsymbol{N}'))) \arrow[d, "\sigma"] \\
 & & H^0(\Omega(\mathrm{Hom}(\boldsymbol{T}, \boldsymbol{T}'))) \arrow[r, "\beta"] \arrow[d, "\gamma"] & H^0(\Omega(\mathrm{Hom}(\boldsymbol{T}, \boldsymbol{N}'))) \arrow[d, "\gamma"] \\
0 \arrow[r] & H^0(\Omega(\mathrm{Hom}(\boldsymbol{T}(X), \boldsymbol{T}(X)))) \arrow[r, "\alpha"] & H^0(\Omega(\mathrm{Hom}(\boldsymbol{T}(X), \boldsymbol{T}'))) \arrow[r, "\beta"] \arrow[d, "\rho"] & H^0(\Omega(\mathrm{Hom}(\boldsymbol{T}(X), \boldsymbol{N}'))) \\
 & & H^1(\Omega(\mathrm{Hom}(\boldsymbol{N}, \boldsymbol{T}'))).
\end{tikzcd}$$

The obstruction to finding $\Lambda \in H^0(\Omega(\mathrm{Hom}(\boldsymbol{T}, \boldsymbol{T}')))$ such that $\gamma(\Lambda) = \alpha(\Psi_*)$ in (1.8) is given by $\rho\alpha(\Psi_*) \in H^1(\Omega(\mathrm{Hom}(\boldsymbol{N}, \boldsymbol{T}')))$. If this obstruction vanishes, then we may find Λ with $\beta(\Lambda) = \sigma(\Psi)$ since $\gamma\beta(\Lambda) = \beta\gamma(\Lambda) = \beta\alpha(\Psi_*) = 0$. Any such Λ is unique if $H^0(\Omega(\mathrm{Hom}(\boldsymbol{N}, \boldsymbol{T}'))) = 0$.

This suffices to check Proposition 1.3 at the first level.

II. The Normal Bundle and the Extension Problem.

1. Some implications of the extension and uniqueness problems for cohomology. Let $X \subset W$ represent a germ of embedding, and let $n = \dim X$, $n + r = \dim W$. For non-negative integers s, t we introduce the following condition:

$*(s,t)$: For any holomorphic vector bundle $\boldsymbol{F}^{\#} \to W$, the stable extension problem for $H^q(X, \Omega(\boldsymbol{F}))$ has a solution for $0 \leqq q < s - 1$, $n - t < q$. Furthermore, we have stable uniqueness for $0 \leqq q < s$, $n - t + 1 < q$.

Remark. If $*(s,t)$ is satisfied, then

$$\dim H^q(W, \Omega(\boldsymbol{F}^{\#})) < \infty \text{ for } 0 \leqq q < s,\ q > n - t + 1. \tag{2.1}$$

PROPOSITION 2.1. *Assume $*(s,t)$ is satisfied. Then there exists a μ_0 such that*

$$H^q(X, \Omega(\boldsymbol{F}(\mu))) = 0 \ \textit{for}\ 0 \leqq q < s - 1,\quad q > n - t + 1, \mu \geqq \mu_0. \tag{2.2}$$

Proof. To say that the stable extension problem has a solution means that, for $\mu \geqq \mu_0(\boldsymbol{F})$, we have

$$H^q(W, \Omega(\boldsymbol{F}^{\#})) \to H^q(\mathfrak{O}^{(\mu)}(\boldsymbol{F}^{\#})) \to 0; \tag{2.3}$$

stable uniqueness means that, for $\mu \geqq \mu_0$, we have

$$0 \to H^q(W, \Omega(\boldsymbol{F}^{\#})) \to H^q(\mathfrak{O}^{(\mu)}(\boldsymbol{F}^{\#})). \tag{2.4}$$

From the exact sheaf diagram

$$\begin{tikzcd}
 & 0 \arrow[d] & \\
 & \Omega(\boldsymbol{F}(\mu+1)) \arrow[d] & \\
\Omega(\boldsymbol{F}^{\#}) \arrow[r] \arrow[d, equal] & \mathfrak{O}^{(\mu+1)}(\boldsymbol{F}^{\#}) \arrow[r] \arrow[d] & 0 \\
\Omega(\boldsymbol{F}^{\#}) \arrow[r] & \mathfrak{O}^{(\mu)}(\mathbf{F}^{\#}) \arrow[r] \arrow[d] & 0 \\
 & 0, &
\end{tikzcd} \tag{2.5}$$

we get a cohomology diagram

$$\begin{tikzcd}
0 \arrow[r] & H^q(\Omega(\boldsymbol{F}^{\#})) \arrow[r] \arrow[d, equal] & H^q(\mathfrak{D}^{(\mu+1)}(\boldsymbol{F}^{\#})) \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & H^q(\Omega(\boldsymbol{F}^{\#})) \arrow[r] & H^q(\mathfrak{D}^{(\mu)}(\boldsymbol{F}^{\#})) \arrow[r] \arrow[d] & 0 \\
 & & H^{q+1}(\Omega(\boldsymbol{F}(\mu+1))) \arrow[d] & \\
0 \arrow[r] & H^{q+1}(\Omega(\boldsymbol{F}^{\#})) \arrow[r] \arrow[d, equal] & H^{q+1}(\mathfrak{D}^{(\mu+1)}(\boldsymbol{F}^{\#})) \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & H^{q+1}(\Omega(\boldsymbol{F}^{\#})) \arrow[r] & H^{q+1}(\mathfrak{D}^{(\mu)}(\boldsymbol{F}^{\#})) \arrow[r] & 0
\end{tikzcd} \tag{2.6}$$

which, by $*(s,t)$, (2.3), and (2.4), is valid for $0 \leqq q+1 < s-1$, $q > n-t+1$. It follows that $H^q(X, \Omega(\boldsymbol{F}(\mu))) = 0$ for $\mu \geqq \mu_0$ and $0 \leqq q < s-1$, $q > n-t+1$.

PROPOSITION 2.2. *Assume that there exists* $\mu_0(\boldsymbol{F})$ *such that*

$$H^q(X, \Omega(\boldsymbol{F}(\mu))) = 0 \text{ for } \mu \geqq \mu_0,\ 0 \leqq q < s,\ q > n-t. \tag{2.7}$$

Then $*(s,t)$ *holds formally.*

Proof. This follows from the exact sequence

$$H^q(\Omega(\boldsymbol{F}(\mu+1))) \to H^q(\mathfrak{D}^{(\mu+1)}(\boldsymbol{F}^{\#})) \to H^q(\mathfrak{D}^{(\mu)}(\boldsymbol{F}^{\#})) \to H^{q+1}(\Omega(\boldsymbol{F}(\mu+1)))$$

since the formal stable extension problem means that

$$H^q(\mathfrak{D}^{(\mu+1)}(\boldsymbol{F}^{\#})) \to H^q(\mathfrak{D}^{(\mu)}(\boldsymbol{F}^{\#})) \to 0 \qquad (\mu \geqq \mu_0), \tag{2.8}$$

whereas formal stable uniqueness means that

$$0 \to H^q(\mathfrak{D}^{(\mu+1)}(\boldsymbol{F}^{\#})) \to H^q(\mathfrak{D}^{(\mu)}(\boldsymbol{F}^{\#})) \qquad (\mu \geqq \mu_0). \tag{2.9}$$

Remark. Conditions (2.1) are by no means independent of those of the form (2.2). For example: Let $\boldsymbol{E} \xrightarrow{\pi} X$ be a holomorphic vector bundle and embed $X \subset \boldsymbol{E}$ as zero cross-section. Let W be a small tubular neighborhood of $X \subset \boldsymbol{E}$ and let $\boldsymbol{F}^{\#} = \pi^{-1}(\boldsymbol{F}) \mid W$ for any holomorphic bundle $\boldsymbol{F} \to X$. We observe here that the normal bundle of $X \subset W$ is just $\boldsymbol{E}$.

LEMMA 2.1. $\sum_{\mu=0}^{\infty} \dim H^q(X, \Omega(\boldsymbol{F}(\mu))) \leqq \dim H^q(W, \Omega(\boldsymbol{F}^{\#}))$.

This is easy to verify using local coordinates. We may in fact sharpen it to the following:

$$H^q(\mathfrak{D}^{(\mu)}(\boldsymbol{F}^{\#})) \cong \bigoplus_{0 \leqq \nu \leqq \mu} H^q(X, \Omega(\boldsymbol{F}(\nu))). \tag{2.10}$$

The content of this section may be summarized as follows: The solution of the stable extension and uniqueness problems for $X \subset W$ imply: (i) a *vanishing theorem* involving the normal bundle $\boldsymbol{N}$ of $X \subset W$, and (ii) a *finite dimensionality theorem* for cohomology groups $H^q(W, \Omega(\boldsymbol{F}^{\#}))$. Conversely, (i) gives at least the formal stable extension and uniqueness problems. Part of our program will be to formulate (i) in such a fashion that it gives the actual extension and uniqueness problems (and thence (ii)).

2. Positive and negative holomorphic bundles. It is well-known what it means for a line bundle $\boldsymbol{E} \rightarrow X$ to be positive, and recently there have been some generalizations of this to general holomorphic bundles (cf. [6] and [8]). For our purposes, we shall need the two notions of *weak positivity* and *positivity* for a holomorphic bundle $\boldsymbol{E} \rightarrow X$; the precise definitions are given in VI. 5 below and we shall only list some relevant properties here: (i) $\boldsymbol{E} \rightarrow X$ is *weakly negative* (*negative*) if, and only if, $\boldsymbol{E}^* \rightarrow X$ is weakly positive (positive); (ii) if $\boldsymbol{E} \rightarrow X$ is weakly positive, then, for some μ_0, the bundles $\boldsymbol{E}^{\mu} (\mu \geqq \mu_0)$ are positive; (iii) for line bundles, weak positivity $=$ positivity; and (iv) if $\boldsymbol{E} \rightarrow X$ is weakly positive and if $\boldsymbol{F} \rightarrow X$ is a bundle, then

$$(2.11) \qquad H^q(X, \Omega(\boldsymbol{F} \otimes \boldsymbol{E}^{\mu})) = 0 \text{ for } \mu \geqq \mu_0(\boldsymbol{F}), \qquad q > 0.$$

More generally, given positive integers s, t with $s + t \leqq n$, we shall define (VI. 5) what it means for $\boldsymbol{E} \rightarrow X$ to have *signature* (s, t); the main properties are: (i) if $\boldsymbol{E} \rightarrow X$ has signature (s, t), then $\boldsymbol{E}^* \rightarrow X$ has signature (t, s); (ii) signature $(n, 0) \equiv$ weak positivity; and (iii) if $\boldsymbol{E} \rightarrow X$ has signature (s, t) and if $\boldsymbol{F} \rightarrow X$ is any bundle, then

$$(2.12) \qquad H^q(X, \Omega(\boldsymbol{F} \otimes \boldsymbol{E}^{*\mu})) = 0 \text{ for } \mu \geqq \mu_0(\boldsymbol{F}), \ 0 \leqq q < s$$

and $q > n - t$.

An example might clarify what it means for $\boldsymbol{E} \rightarrow X$ to have signature (s, t). Let Y, Z be algebraic manifolds of dimensions s, t respectively, let $\boldsymbol{F} \rightarrow Y$ be a positive line bundle and $\boldsymbol{G} \rightarrow Z$ a negative line bundle. Then $\boldsymbol{E} = \boldsymbol{F} \otimes \boldsymbol{G}$ has signature (s, t) over $X = Y \times Z$. Observe that, for large μ, $H^q(Y, \Omega(\boldsymbol{F}^{*\mu})) = 0$ for $q < s$ and $H^q(Z, \Omega(\boldsymbol{G}^{*\mu})) = 0$ for $q > 0$. By Künneth, $H^q(X, \Omega(\boldsymbol{E}^{*\mu})) = 0$ for $0 \leqq q < s$ and $q > n - t = s$.

From 2.12 and Propositions 1.1-1.4, we have

PROPOSITION 2.3. *Let $X \subset W$ define a germ of embedding whose normal bundle $\boldsymbol{N}$ has signature (s, t). Then:*

(i) *The statement $*(s, t)$ in § II. 1 holds formally;*

(ii) *The rigidity problem is formally true provided that either* $s \geqq 2$ *or* $t = n$. *In the former case, we have uniqueness;*

(iii) *The extension problem for holomorphic bundles is formally true if either* $s \geqq 3$ *or* $t \geqq n-1$. *In the former case we have formal uniqueness always, and we have formal uniqueness in the latter case if* $t = n$.

Assume now that the normal bundle $\boldsymbol{N}$ of $X \subset W$ has signature (s,t) and let $\boldsymbol{F}^{\#} \to W$ be a holomorphic vector bundle. Set $\Omega_\mu(\boldsymbol{F}^{\#}) = \mathfrak{J}^{\mu+1} \cdot \Omega(\boldsymbol{F}^{\#})$ so that we have

$$0 \to \Omega_\mu(\boldsymbol{F}^{\#}) \to \Omega(\boldsymbol{F}^{\#}) \to \mathfrak{O}^{(\mu)}(\boldsymbol{F}^{\#}) \to 0; \tag{2.13}$$

$$0 \to \Omega_{\mu+\nu}(\boldsymbol{F}^{\#}) \to \Omega_\mu(\boldsymbol{F}^{\#}) \quad (\nu \geqq 0); \text{ and} \tag{2.14}$$

$$0 \to \Omega_{\mu+1}(\boldsymbol{F}^{\#}) \to \Omega_\mu(\boldsymbol{F}^{\#}) \to \Omega(\boldsymbol{F}(\mu+1)) \to 0. \tag{2.15}$$

Proposition 2.4. *For* $\mu \geqq \mu_0$ *and all* $\nu \geqq 0$ *the sequence* (2.14) *induces an isomorphism*

$$0 \to H^q(W, \Omega_{\mu+\nu}(\boldsymbol{F}^{\#})) \to H^q(W, \Omega_\mu(\boldsymbol{F}^{\#})) \to 0 \tag{2.16}$$

for $0 \leqq q < s,\ q > n - t + 1$.

Proof. This follows inductively on ν from the exact cohomology sequence of (2.15) together with (2.12).

Proposition 2.5. *If, for large* μ,

$$H^q(W, \Omega_\mu(\boldsymbol{F}^{\#})) = 0 \text{ for } 0 \leqq q < s,\ q > n - t + 1,$$

then $*(s,t)$ *holds.*

Proof. This follows from (2.3), (2.4), and the exact cohomology sequence of (2.13).

Definition. If the conclusion (2.16) of Proposition 2.4 holds, then we shall say that the group $H^q(W, \Omega_\mu(\boldsymbol{F}^{\#}))$ *vanishes to infinite order* along X.

By Proposition 2.5, the *actual* existence theorem for the extension and uniqueness problems in cohomology reduces to showing that cohomology which vanishes to infinite order along X is necessarily zero. This seems to be a standard procedure for deducing existence theorems from formal statements.

3. Statement of the main results.

Theorem I. *Let* $X \subset W$ *define a germ of embedding whose normal bundle* $\boldsymbol{N}$ *has signature* (s,t) *with* $s + t = n$. *Then, for any holomorphic* *vector bundle* $\boldsymbol{F}^{\#} \rightarrow W$, *the stable extension problem for* $H^q(X, \Omega(\boldsymbol{F}))$ *has a solution for* $0 \leqq q < s-1$, $q > n-t$. *Furthermore, stable uniqueness holds for* $0 \leqq q < s$, $q > n-t+1$.

We shall use Theorem I to discuss the rigidity question and extension problem for holomorphic bundles in certain special cases. In order to state the results, we need the notion of a sufficiently positive bundle.

Definition. Let $\boldsymbol{E} \rightarrow X$ be a positive holomorphic vector bundle. Associated to X there is a finite number $\boldsymbol{F}_1, \cdots, \boldsymbol{F}_k$ of holomorphic bundles, and $\boldsymbol{E}$ is defined to be *sufficiently positive* if all the bundles $\boldsymbol{E} \otimes \boldsymbol{F}_i$ $(1 \leqq i \leqq k)$ are positive.

Remark. For a positive (but *not* in general weakly positive) bundle $\boldsymbol{E} \rightarrow X$, we have

$$H^q(X, \Omega(\boldsymbol{E}^{*\mu})) = 0 \qquad \text{for } q > 0, \mu > 0. \tag{2.17}$$

THEOREM II. *Let* $X \subset W$ *define a germ of embedding whose normal bundle* $\boldsymbol{N} \rightarrow X$ *is sufficiently positive.*

(i) *The germ of embedding is rigid provided that* $\dim X \geqq 3$.

(ii) *The extension problem for holomorphic bundles holds of* $\dim X \geqq 3$.

(iii) *Stable uniqueness holds in* (i) *and* (ii).

Remark. The hypothesis of "sufficient positivity" should be thought of as describing a class $\mathcal{S}$ of holomorphic vector bundles over X such that Theorem II is true for $X \subset W$ whenever $\boldsymbol{N} \in \mathcal{S}$. For codimension 1, $\mathcal{S} =$ all positive line bundles over X; this is simply because passing from the linear system $|\boldsymbol{L}|$ to $|\mu\boldsymbol{L}|$ $(\boldsymbol{L} = [X])$ corresponds to replacing $\boldsymbol{N} \rightarrow X$ by $\boldsymbol{N}^{\mu}$ which will be sufficiently positive for μ large. For codimension $r > 1$, there is no such (obvious) way of making a subvariety more positive and this is basically the reason for the hypothesis of sufficient positivity.

We ashall prove Theorem I by proving

THEOREM III. *With the notations and assumptions of Theorem* I, *we have*

$$H^q(W, \Omega_{\mu}(\boldsymbol{F}^{\#})) = 0 \text{ for } \mu \geqq \mu_0,\ 0 \leqq q < s \text{ or } q > n-t+1. \tag{2.18}$$

We now observe

PROPOSITION 2.6. *Let* $X \subset W$ *have positive normal bundle* $\boldsymbol{N}$ *and assume that* $\dim X \geqq 3$. *Then, assuming* (2.18), *the extension problem for a line bundle* $\boldsymbol{E} \rightarrow X$ *has unique solution.*

Proof. We shall verify (1.5). The middle line there is clear, and we have the exact sequence

$$H^1(W,\mathfrak{O}_W)\to H^1(X,\mathfrak{O}_X)\to H^2(W,\mathfrak{J})\to H^2(W,\mathfrak{O}_W)\to H^2(X,\mathfrak{O}_X).$$

It will suffice to prove that $H^2(W,\mathfrak{J})=0$. From the sequences

$$0\to\mathfrak{J}^{\mu+1}\to\mathfrak{J}^{\mu}\to\Omega(\mu)\to 0,$$

we get $0\to H^2(W,\mathfrak{J}^{\mu+1})\to H^2(W,\mathfrak{J}^{\mu})\to 0$ since $H^q(X,\Omega(\mu))=0$ for $\mu\geqq 1$, $q\leqq 2\leqq \dim X-1$. This then gives $0\to H^2(W,\mathfrak{J}^{\mu})\to H^2(W,\mathfrak{J})\to 0$ for all $\mu\geqq 1$ which, by (2.18), gives $H^2(W,\mathfrak{J})=0$, Q. E. D.

4. Outline of the proof of Theorem III. We give now an outline of the proof of Theorem III; the details and relevant calculations will be supplied in §§ V-VIII below. Thus let $X\subset W$ be an embedding of codimension r and where the normal bundle $\boldsymbol{N}$ has signature (s,t) with $s+t=n=\dim X$. By blowing up $X\subset W$, we obtain a hypersurface $\tilde{X}\subset\tilde{W}$ (§ V.1) and the normal bundle $\tilde{\boldsymbol{N}}$ of $\tilde{X}\subset\tilde{W}$ will have signature $(s,t+r-1)$ (VI.7). Furthermore, it is at least plausible that, in order to examine cohomology vanishing on $X\subset W$, we may lift everything up and look at cohomology vanishing on $\tilde{X}\subset\tilde{W}$ (V.3).

We remark here that, even if $X\subset W$ has a positive normal bundle, then the normal bundle of $\tilde{X}\subset\tilde{W}$ will be of mixed signature, and this is the essential difficulty in the higher codimension case. In fact, if $X\subset W$ is a hypersurface whose normal bundle has either signature $(n,0)$ or $(0,n)$, then Theorem III is easy to prove directly (IV.1 and [6], § 4).

Thus we may assume that $r=1$. In this case, the sheaf

$$(2.19)\qquad \Omega_\mu(\boldsymbol{F}^\#)=\Omega(\boldsymbol{F}^\#\otimes\boldsymbol{L}^{*\mu})=\Omega(\boldsymbol{F}^\#(\mu))\qquad(\boldsymbol{L}=[X])$$

is locally free and we may hope to use the theory of harmonic integrals on open manifolds as developed by J. Kohn (VIII.1 and [15]) to study the groups $H^q(W,\Omega(\boldsymbol{F}^\#(\mu)))$.

Now in order to apply this harmonic integral theory, we must have an exhaustion $W=\bigcup_{0<c<c_0}W_c$ of W by nice relatively compact subdomains $W_c=\{w\in W\colon r(w)<c\}$ for some non-negative function r on W. Furthermore, for some $\epsilon>0$, $\epsilon<c_0$, $\partial\bar{\partial}r$ in $W-W_\epsilon$ should have signature $(t+1,s)$. Now, since $\boldsymbol{N}$ has signature (s,t) with $s+t=n$, such an r exists for the embedding $X\subset\boldsymbol{N}$ as zero cross-section (VI.5) and, by an approximation argument, r exists for $X\subset W$ provided that W is small enough (VIII.3).

If now $\partial\bar{\partial}r$ had signature $(t+1,s)$ everywhere in W, then the trivial bundle $\boldsymbol{l} \to W$ would have signature $(t+1, s)$ and it is fairly easy to conclude then that (VII. 4)

$$(2.20) \qquad H^q(W, \Omega(\boldsymbol{E}^\#)) = 0 \quad \text{for } 0 \leqq q < s, q > n - t$$

for *every* holomorphic bundle $\boldsymbol{E}^\# \to W$. This is the so-called complete case and has been discussed in [2].

Now our function r is certainly *not* complete since $\partial\bar{\partial}r \mid X = \partial\bar{\partial}(r \mid X) = 0$. Thus (2.20) will not hold in general—all we can conclude is that

$$(2.21) \qquad \dim H^q(W, \Omega(\boldsymbol{E}^\#)) < \infty \qquad 0 \leqq q < s, q > n - t.$$

However, it is perhaps not unreasonable to expect that (2.20) will hold provided that we can find in $\boldsymbol{E}^\# \to W$ a metric whose curvature form picks up the slack caused by the degeneracy of $\partial\bar{\partial}r$. In fact, this is the whole point: For μ sufficiently large, $\boldsymbol{F}^\#(\mu) = \boldsymbol{F}^\# \otimes \boldsymbol{L}^{*\mu}$ will have signature $(t, s+1)$ (VIII. 1) and furthermore the curvature form in $\boldsymbol{F}^\#(\mu)$ will have an s-dimensional negative eigenspace close to the s-dimensional negative eigenspace of $\partial\bar{\partial}r$ on $W - W_\epsilon$, and similarly for the t-dimensional positive spaces. (Observe here that there is a loss of 1-dimensional between (2.18) and (2.20).)

Thus, if one accepts this argument (which is almost obvious geometrically) and combines it with inequalities obtained from two integral formulae on open manifolds (VII. 3 and VII. 6), then Theorem III follows easily from the Bochner vanishing technique together with Kohn's results about representing cohomology by harmonic forms with boundary conditions.

III. Positive Embeddings and the Proof of Theorem II.

1. A simple example of rigidity. We shall prove a simple example of the rigidity theorem which will, however, contain the main ideas of the general argument to be given below. Let $X = P_n$ $(n \geqq 3)$ and let $\boldsymbol{H} \to X$ be the hyperplane bundle; $\boldsymbol{H} = [h]$ where $h \subset P_n$ is a hyperplane. Suppose now that we have an embedding $X \subset W$ with normal bundle $\boldsymbol{H}$.

LEMMA 3.1. $H^1(X, \Omega(\boldsymbol{T}(\mu))) = 0$ *for* $\mu > 0$.
The proof of this will be given in IV. 2. (i) below.

Thus we see that the germ of embedding defined by $X \subset W$ should be equivalent to the germ defined by embedding $X \subset \boldsymbol{H}$ as zero cross-section, and we shall use Theorem I to prove that this is the case. Note that this example is "dual" to the rigidity of quadratic transformations (cf. the appendix to [13]).

Set $\boldsymbol{L} = [X]$ so that $\mathfrak{J} \cong \Omega(\boldsymbol{L}^*)$ and $\boldsymbol{L} \mid X = \boldsymbol{H}$. We have the exact sheaf sequence

$$0 \to \mathfrak{O}_W \to \Omega(\boldsymbol{L}) \to \Omega_X(\boldsymbol{H}) \to 0 \tag{3.1}$$

LEMMA 3.2. $H^1(W, \mathfrak{O}_W) = 0$.

Proof. Since $H^1(X, \mathfrak{O}_X) = 0$, this will follow from $H^1(W, \mathfrak{J}) = 0$. But $H^1(W, \mathfrak{J}) = 0$ by Theorem I and since $\dim X \geqq 3$ (cf. the proof of Proposition 2.6). Q.E.D.

From (3.1) we get then

$$0 \to \boldsymbol{C} \to H^0(W, \Omega(\boldsymbol{L})) \xrightarrow{\rho} H^0(X, \Omega(\boldsymbol{H})) \to 0. \tag{3.2}$$

Thus $\dim H^0(W, \Omega(\boldsymbol{L})) = n + 2$ and we choose a basis $\xi_0, \cdots, \xi_{n+1}$ of $H^0(W, \Omega(\boldsymbol{L}))$ so that $\rho(\xi_{n+1}) = 0$. If we now define $\Phi\colon W \to P_{n+1}$ by

$$\Phi(w) = [\xi_0(w), \cdots, \xi_n(w), \xi_{n+1}(w)],$$

then an elementary consideration using the complete (since $H^1(X, \Omega(\boldsymbol{H})) = 0$) continuous system of $X \subset W$ (cf. III.2 and III.3) shows that Φ is an embedding and $\Phi(X) = P_n \subset P_{n+1}$. Thus $X \subset W$ defines the same germ as the usual embedding $P_n \subset P_{n+1}$.

On the other hand, let $S = \{\xi \in P_{n+1} : \xi_0 = \cdots = \xi_n = 0\}$. Then $S = \text{pt.}$ and $P_{n+1} - S$ is a neighborhood W' of $P_n \subset P_{n+1}$. Furthermore, we may define $\pi\colon W' \to P_n$ by $\pi[\xi_0, \cdots, \xi_n, \xi_{n+1}] = [\xi_0, \cdots, \xi_n]$ and then W' is a line bundle over P_n, and in fact $W' \cong \boldsymbol{H}$ as line bundles over P_n. Thus $X \subset W$ and $X \subset \boldsymbol{H}$ define equivalent germs of embeddings.

Remark. The compactification of $\boldsymbol{H}$ by adding on S to obtain a closed variety is a general phenomenon (cf. IV.1).

2. On the theory of continuous systems. Grauert has shown, if the normal bundle $\boldsymbol{N}$ of X in W is negative, then we may assume that W is *strongly-pseudo-convex*, that $X \subset W$ is the only non-trivial compact subvariety, and that there exists a Stein space Y and a proper holomorphic mapping $\pi\colon W \to Y$ such that $\pi(X) = y_0$ is a point and $\pi\colon W - X \cong Y - \{y_0\}$.

Now in case $\boldsymbol{N}$ is positive it is again true that we may assume that W is *strongly pseudo-concave*,* but we need better information than this alone. Namely, we shall need the fact that, if $\boldsymbol{N}$ is sufficiently positive, then the *characteristic system* of X in W is complete and we may write $W = \{X_t\}_{t \in B}$ where the X_t form an analytic family of compact submanifolds of W which are small deformations of X in W.

---

* Cf. VIII 3.

To be more specific, let $X \subset W$ be a germ of embedding with normal bundle $\boldsymbol{N}$. Let $\rho_x \colon H^0(X,\Omega(\boldsymbol{N})) \to \boldsymbol{N}_x$ be the restriction mapping and set $\delta(x) = \dim \rho_x\{H^0(X,\Omega(\boldsymbol{N}))\}$. Let $B \subset \boldsymbol{C}^m$ be an open neighborhood of the origin with linear coordinates $t = (t^1,\cdots,t^m)$. In [12] Kodaira has defined the concept of an *analytic family* $\{\boldsymbol{V},B\}$ *of compact submanifolds* of W parametrized by B. We write $\boldsymbol{V} = \{X_t\}_{t\in B}$ with $X_0 = X$. By a *deformation* $\{\boldsymbol{V} \xrightarrow{\tilde{\omega}} B\}$ of X, with parameter space B, we mean a complex manifold $\boldsymbol{V}$ together with an onto projection $\tilde{\omega}\colon \boldsymbol{V} \to B$ which is everywhere of maximal rank and where $\tilde{\omega}^{-1}(0) \cong X$.

THEOREM 3.1. *Assume* $H^1(X,\Omega(\boldsymbol{N})) = 0$. *Then there exists*:

(i) *a neighborhood* B *of the origin in* $H^0(X,\Omega(\boldsymbol{N}))$;

(ii) *a family* $\mathfrak{V} = \{X_t\}_{t\in B}$ *of compact analytic submanifolds of* W;

(iii) *a deformation* $\{\boldsymbol{V} \xrightarrow{\tilde{\omega}} B\}$; *and*

(iv) *a holomorphic mapping* $F\colon \boldsymbol{V} \to W$ *such that* $F(\tilde{\omega}^{-1}(t)) = X_t$ *and such that rank of* F *at* $x \in X \subset V$ *is* $\delta(x) + \dim X$.

Proof. This result will be a consequence of the main theorem in [12]; we shall use the proof as an opportunity to establish some notation for later use. Let $\dim X = n$, $\dim W = n + r$, $\dim H^0(X,\Omega(\boldsymbol{N})) = m$, and agree on the ranges of indices $1 \leqq i,j \leqq n$; $1 \leqq a,b \leqq r$; and $1 \leqq \rho,\sigma \leqq m$. We choose a covering $\{U_\alpha\}$ of W by open sets such that, if $V_\alpha = U_\alpha \cap X$, then $U_\alpha \cap U_\beta = \emptyset \Longleftrightarrow V_\alpha \cap V_\beta = \emptyset$. We assume that we have in U_α holomorphic coordinates $(z_\alpha, w_\alpha) = (z_\alpha{}^1,\cdots,z_\alpha{}^n; w_\alpha{}^1,\cdots,w_\alpha{}^r)$ such that $U_\alpha \cap X$ is given by $\{w_\alpha = 0\}$. If we set $|z_\alpha| \max_{1\leqq i\leqq n} |z_\alpha{}^i|, |w_\alpha| \max_{1\leqq a\leqq r} |w_\alpha{}^a|$, we may suppose that U_α is the unit polycylinder $|z_\alpha| < 1$, $|w_\alpha| < 1$.

In $U_\alpha \cap U_\beta$ we have

$$\begin{cases} z_\alpha = f_{\alpha\beta}(z_\beta, w_\beta) \\ w_\alpha = g_{\alpha\beta}(z_\beta, w_\beta)\,;\, g_{\alpha\beta}(z_\beta, 0) = 0. \end{cases}$$

The normal bundle $\boldsymbol{N}$ has transition functions given by the $r \times r$ matrix $N_{\alpha\beta}(z_\beta) = \{N_{\alpha\beta}(z_\beta)_b{}^a\}$ where

$$N_{\alpha\beta}(z_\beta)_b{}^a = \frac{\partial g_{\alpha\beta}{}^a(z_\beta, w_\beta)}{\partial w_\beta{}^b} \Big\}_{w=0}$$

We choose a basis $\sigma_1,\cdots,\sigma_m$ of $H^0(X,\Omega(\boldsymbol{N}))$. Then $\sigma_\rho \mid U_\alpha = \sigma_{\rho,\alpha}$ is column vector ${}^t(\sigma_{\rho,\alpha}{}^1,\cdots,\sigma_{\rho,\alpha}{}^r) = {}^t\{\sigma_{\rho,\alpha}{}^a\}$ * and satisfies $\sigma_{\rho,\alpha} = N_{\alpha\beta}\sigma_{\rho,\beta}$ in

---

* ${}^t(\#)$ means transpose of the row vector $(\#)$.

$V_\alpha \cap V_\beta$. A general point in $H^0(X, \Omega(N)) \cong \boldsymbol{C}^m$ will be written as $t = \sum_{\rho=1}^{m} t^\rho \sigma_\rho$. By Kodaira's theorem [12], there exists a neighborhood B of the origin in $\boldsymbol{C}^m$ and a family of vector-valued holomorphic functions $\phi_\alpha(z_\alpha, t) : V_\alpha \times B \rightarrow \boldsymbol{C}^r$ which satisfy the following conditions:

$$\phi_\alpha(z_\alpha, 0) = 0 \text{ and } |\phi_\alpha(z_\alpha, t)| < 1; \tag{3.3}$$

$$\left.\frac{\partial \phi_\alpha(z_\alpha, t)}{\partial t^\rho}\right]_{t=0} = \sigma_{\rho,\alpha}(z_\alpha); \tag{3.4}$$

$$\phi_\alpha(z_\alpha, t) = g_{\alpha\beta}(z_\beta, \phi_\beta(z_\beta, t)); \text{ and} \tag{3.5}$$

$$z_\alpha = f_{\alpha\beta}(z_\beta, \phi_\beta(z_{\beta,t})). \tag{3.6}$$

If we define $X_t \subset W$ by $X_t \cap U_\alpha = \{(z_\alpha, w_\alpha) : w_\alpha = \phi_\alpha(z_\alpha, t)\}$, then by (3.3), (3.5), and (3.6) X_t is a compact sub-manifold of W and $X_0 = X$. By (3.4), we may symbolically write $\left.\frac{\partial X_t}{\partial t^\rho}\right]_{t=0} = \sigma_\rho$. The collection $\{X_t\}_{t \in B}$ clearly gives an analytic family $\mathfrak{B}$ of compact submanifolds of W.

We construct an abstract complex manifold $\boldsymbol{V}$ as follows: Let $T_\alpha = V_\alpha \times B$. Then $\boldsymbol{V}$ is formed from the union $\bigcup_\alpha T_\alpha$ by identifying (z_α, t) with (z_β, t') if, and only if, $t = t'$ and $z_\alpha = f_{\alpha\beta}(z_\beta, \phi_\beta(z_\beta, t))$. From (3.5) and (3.6) it follows that this is an equivalence relation, and so $\boldsymbol{V}$ is a complex manifold. Furthermore, if we define $\tilde{\omega}: \boldsymbol{V} \rightarrow B$ by $\omega(z_\alpha, t) = t$ in T_α, $\tilde{\omega}$ is clearly well-defined and has maximal rank everywhere. Obviously then $\{\boldsymbol{V} \xrightarrow{\tilde{\omega}} B\}$ is a deformation of X.

We define $F: \boldsymbol{V} \rightarrow W$ as follows: $F_\alpha = F \mid T_\alpha$ is given by $F_\alpha(z_\alpha, t) = (z_\alpha, \phi_\alpha(z_\alpha, t)) \in U_\alpha$ by (3.3). It follows from (3.5) and (3.6) that $F_\alpha = F_\beta$ in $T_\alpha \cap T_\beta$, and thus F is globally defined. The Jacobian matrix $J(F_\alpha)$ of F_α, when evaluated on $X \cap T_\alpha$, is given by the $(n+m) \times (n+r)$ matrix $J(F_\alpha(z_\alpha)) = \begin{pmatrix} \boldsymbol{I}^n & * \\ 0 & Q(z_\alpha) \end{pmatrix}$ where $\boldsymbol{I}^n$ is the $n \times n$ unit matrix and $Q(z_\alpha)$ is the $m \times r$ matrix with entries

$$Q(z_\alpha)_\rho{}^a = \left.\frac{\partial \phi_\alpha{}^a(z_\alpha, t)}{\partial t^\rho}\right]_{t=0} = \sigma_{\rho,\alpha}{}^a(z_\alpha). \tag{3.7}$$

For fixed $z_\alpha{}^0 \in V_\alpha$, the vectors $\sigma_{1,\alpha}(z_\alpha{}^0), \cdots, \sigma_{m,\alpha}(z_\alpha{}^0)$ span a subspace of $\boldsymbol{C}^r$ of dimension $\delta(z_\alpha{}^0)$, and thus the rank of $J(F_\alpha(z_\alpha{}^0))$ is $\delta(z_\alpha{}^0)$. Q.E.D.

3. Positive embeddings and continuous systems. From Theorem 3.1 we obtain:

Proposition 3.1. *If $H^1(X,\Omega(N))=0$ and if the restriction mapping $\rho_x\colon H^0(X,\Omega(N))\to N_x$ is everywhere surjective, then $F\colon V\to W$ is onto and we may write $W=\{X_t\}_{t\in B}$.*

Remark. In [7] we assumed that, if N is trivial, then W is a space of deformations of X. This is not so, as the example below shows, but we do have:

Proposition. *If $H^1(X,\mathfrak{O})=0$ and if N is trivial, then W is a space of deformations of X.*

Proof. $\Omega(N)\cong\underbrace{\mathfrak{O}\oplus\cdots\oplus\mathfrak{O}}_{r}$ since N is trivial. Then $H^1(X,\Omega(N))=0$ and $\rho_x\colon H^0(X,\Omega(N))\to N_x$ is onto for all x. Consequently, $F\colon V\to W$ is biregular in a neighborhood of X and so W is a space of deformations of X.

Example. We show that, if $H^1(X,\mathfrak{O})\neq0$, then there exists an embedding $X\subset W$ in which N is trivial but where no neighborhood of X in W is a deformation space of X. Let $h_{\alpha\beta}\colon V_\alpha\cap V_\beta\to C$ be a 1-cocycle giving a non-zero element in $H^1(X,\mathfrak{O})$. For each complex number t, $|t|<\epsilon$, define a new complex manifold W_t from the union $\bigcup_\alpha\{V_\alpha\times C\}$ by identifying (z_α,ξ_α) with (z_β,ξ_β) if, and only if, $z_\alpha=z_\beta$ and $\xi_\alpha=\exp(th_{\alpha\beta}(z_\beta))\cdot\xi_\beta$. Then W_t is a line bundle over X which, for ϵ small, has a non-identically zero cross-section if, and only if, $t=0$. Let $W=\bigcup_{|t|<\epsilon}W_t$ and embed $X\subset W_0\subset W$ as zero cross-section. Clearly, no neighborhood of X in W is a deformation space of X. On the other hand, by (6.2) the normal bundle N has transition functions

$$\left.\begin{pmatrix}\dfrac{\partial\xi_\alpha}{\partial\xi_\beta} & \dfrac{\partial\xi_\alpha}{\partial t}\\[2ex] \dfrac{\partial t}{\partial\xi_\beta} & \dfrac{\partial t}{\partial t}\end{pmatrix}\right|_{t=0=\xi}=\begin{pmatrix}1&0\\0&1\end{pmatrix}.$$

Remark. This example is similar to Zappa's example of an incomplete continuous system. Let X, $\{h_{\alpha\beta}\}\in H^1(X,\mathfrak{O})$ be as above, and let $E\to X$ be the bundle with fibre C^2 and transition functions $g_{\alpha\beta}=\begin{pmatrix}1&h_{\alpha\beta}\\0&1\end{pmatrix}$. If $P(E)\xrightarrow{\pi}X$ is the associated projective bundle, then $P(E)$ has a covering by open sets $U_\alpha=\pi^{-1}(V_\alpha)$ with coordinates $(z_\alpha;[\xi_\alpha{}^0,\xi_\alpha{}^1])$ in U_α and the transition rules $\xi_\alpha{}^0=\xi_\beta{}^0+h_{\alpha\beta}(z_\beta)\xi_\beta{}^1$, $\xi_\alpha{}^1=\xi_\beta{}^1$. Thus the cross-section S at infinity given by $S\cap U_\alpha=\{\xi_\alpha{}^1=0\}$ is defined, and the normal bundle

$\boldsymbol{N} \to S$ of $S \subset P(\boldsymbol{E})$ is trivial. Consequently, $H^0(S,\Omega(\boldsymbol{N})) \neq 0$ but the continuous system generated by $S \subset P(\boldsymbol{E})$ is empty. Indeed, if there exists a family $\{S_t\}$ with $S_0 = S$, then $S_t \cap S = \emptyset$ for small t (since $\boldsymbol{N}$ is trivial) and $\pi\colon S_t \to X$ is biregular. Consequently, there exists $\sigma\colon X \to P(\boldsymbol{E})$ with $\sigma(X) \cap S = \emptyset$. If $\sigma \mid V_\alpha = [\sigma_\alpha{}^0(z), \sigma_\alpha{}^1(z)]$, then $\tau_\alpha(z) = \sigma_\alpha{}^0(z)\sigma_\alpha{}^1(z)^{-1}$ satisfies $\tau_\alpha - \tau_\beta = h_{\alpha\beta}$ which is a contradiction.

Now consider the sequences

$$(3.8) \qquad 0 \to \boldsymbol{N}(x,x') \to H^0(X,\Omega(\boldsymbol{N})) \xrightarrow{\rho_{x,x'}} \boldsymbol{N}_x \oplus \boldsymbol{N}_{x'} \qquad (x \neq x')$$

$$(3.9) \qquad \boldsymbol{N}(x) \xrightarrow{\sigma_x} \boldsymbol{m}_x/\boldsymbol{m}_x{}^2 \otimes \boldsymbol{N}_x$$

where $\boldsymbol{N}(x) = \ker \rho_x\colon H^0(X,\Omega(\boldsymbol{N})) \to \boldsymbol{N}_x$ and $\boldsymbol{m}_x \subset \mathfrak{O}_x$ is the maximal ideal at x.

PROPOSITION 3.2. *Assume* $H^1(X,\Omega(\boldsymbol{N})) = 0$ *and that* $\rho_{x,x'}$ *is everywhere surjective. Then, given* $x \in X$ *and* $w \in W$, *there exists an* $X_t \subset W$ *which passes through* x *and* w.

PROPOSITION 3.3. *Assume* $H^1(X,\Omega(\boldsymbol{N})) = 0$ *and that* σ_x *is everywhere surjective. Then, given* $x \in X$ *and* $\xi \in \boldsymbol{T}_x(W)$, *there exists an* $X_t \subset W$ *such that* $x \in X_t$ *and* $\xi \in \boldsymbol{T}_x(X_t)$.

Both Propositions follow from Theorem 3.1. We observe that, if $\boldsymbol{N} \to X$ is sufficiently positive, then $H^1(X,\Omega(\boldsymbol{N})) = 0$ and both $\rho_{x,x'}$ and σ_x are evereywhere onto (cf. [8]). In fact, the class $\mathscr{S}$ of sufficiently positive bundles over X (cf. the footnote to Theorem II) will consist of those holomorphic bundles $\boldsymbol{N} \to X$ for which $\rho_{x,x'}$ and σ_x are everywhere surjective, for which $\boldsymbol{N} \otimes \boldsymbol{K}^*$ is positive, and for which the bundles $\boldsymbol{L}^* \otimes \boldsymbol{N}^\mu$ $(\mu \geqq 1)$ are all positive for some ample line bundle $\boldsymbol{L} \to X$. An embedding $X \subset W$ for which $\boldsymbol{N} \in \mathscr{S}$ will be called a *positive embedding.*

Let $\boldsymbol{L} \to X$ be an ample line bundle so that $|\boldsymbol{L}|$ gives a projective embedding $\lambda\colon X \to P_s(s+1 = \dim H^0(X,\Omega(\boldsymbol{L})))$.

PROPOSITION 3.4. *If* $X \subset W$ *is a positive embedding and if the extension problems for* $\boldsymbol{L}$ *and* $H^0(X,\Omega(\boldsymbol{L}))$ *can be solved, then* λ *extends to a projective embedding* $\Lambda\colon W \to P_s$.

Proof. Let $\boldsymbol{L}^\#$ be an extension to W of $\boldsymbol{L}$. Since $H^0(W,\Omega_\mu(\boldsymbol{L}^\#)) = 0$ for large μ, $\dim H^0(W,\Omega(\boldsymbol{L}^\#)) = S+1 < \infty$. By assumption we also have $H^0(W,\Omega(\boldsymbol{L}^\#)) \to H^0(X,\Omega(\boldsymbol{L})) \to 0$. Thus λ extends to a mapping $\Lambda\colon W \to P_S$. On the other hand, we clearly have that $\Lambda_t = \Lambda \mid X_t$ is an embedding (for t

small). Let $x \in X$ and $\xi \in \boldsymbol{T}_x(W)$. By Proposition 3.3 there exists X_t such that $x \in X_t$ and $\xi \in \boldsymbol{T}_x(X_t)$. Then $\Lambda_*(\xi) = (\Lambda_t)_*(\xi) \neq 0$, and so Λ is of maximal rank in a neighborhood of X. Q. E. D.

COROLLARY. *Let $X \subset W$ be a positive embedding. Then, assuming Theorem* I, *we have that there exists a projective embedding $\Lambda\colon W \to P_N$, provided that* $\dim X \geqq 3$.

Proof. Let $\boldsymbol{L} \to X$ be an ample line bundle for which $\boldsymbol{L}^* \otimes \boldsymbol{N}^\mu$ is positive for $\mu > 0$. Then $H^2(X, \Omega(\mu)) = 0 = H^1(X, \Omega(L(\mu)))$ for $\mu > 0$. By Theorem I we can then solve the extension problems for $\boldsymbol{L}$ and $H^0(X, \Omega(\boldsymbol{L}))$. Q. E. D.

In case $X \subset W$ is a hypersurface, we need only assume that $\dim X \geqq 2$ since the line bundle $\boldsymbol{L} = [X]$ is already an extension of the positive bundle $\boldsymbol{N} = \boldsymbol{L} \mid X$. In fact, we have

PROPOSITION 3.5 [18]. *Let $\boldsymbol{E}^\# \to W$ be a holomorphic vector bundle and $\boldsymbol{E} = \boldsymbol{E}^\# \mid W$. Then, for $\mu \geqq \mu_0(\boldsymbol{E})$, the mapping*

$$H^0(W, \Omega(\boldsymbol{E}^\# \otimes \boldsymbol{L}^\mu)) \to H^0(X, \Omega(\boldsymbol{E} \otimes \boldsymbol{N}^\mu))$$

is onto.

Proof. Since $\mathfrak{J} \cong \Omega(\boldsymbol{L}^*)$, we have

$$0 \to \Omega(\boldsymbol{E}^\# \otimes \boldsymbol{L}^{\mu-1}) \to \Omega(\boldsymbol{E}^\# \otimes \boldsymbol{L}^\mu) \to \Omega_X(\boldsymbol{E} \otimes \boldsymbol{N}^\mu) \to 0 \tag{3.10}$$

for all μ. We choose ν_0 so that $H^1(X, \Omega(\boldsymbol{E} \otimes \boldsymbol{N}^\mu)) = 0$ for $\mu \geqq \nu_0$. From the exact cohomology sequences of (3.10), we get

$$H^1(W, \Omega(\boldsymbol{E}^\# \otimes \boldsymbol{L}^{\nu_0})) \to H^1(W, \Omega(\boldsymbol{E}^\# \otimes \boldsymbol{L}^{\nu_0+\mu}))) \to 0$$

for all $\mu \geqq 0$. On the other hand, we shall see in VII.1 and VIII.3 that $\dim H^1(W, \Omega(\boldsymbol{E}^\# \otimes \boldsymbol{L}^{\nu_0})) < \infty$. Thus, for large enough μ, the mappings

$$H^1(W, \Omega(\boldsymbol{E}^\# \otimes \boldsymbol{L}^{\mu-1})) \xrightarrow{\approx} H^1(W, \Omega((\boldsymbol{E}^\# \otimes \boldsymbol{L}^\mu))$$

are all isomorphisms. The Proposition now follows from this plus the exact cohomology sequence of (3.10).

4. Proof of Theorem II from Theorem I.

Proof of (i). Let $X \subset W$ be a positive germ of embedding. Then there exists a μ_0 such that, if $X \subset W'$ defines a germ of embedding for which we have an equivalence $f_\mu\colon X_\mu' \to X_\mu$, then there is a unique formal extension $\hat{f}\colon \hat{X}' \to \hat{X}$ of f_μ. We want to prove that $\hat{f}$ converges.

By the Corollary to Proposition 3.4, there exists an ample line bundle $\boldsymbol{L}_* \rightarrow W_*$ which gives a projective embedding $\Lambda : X_* \rightarrow P_S$.

Now we assert that, if we are given *two* projective embeddings $\Lambda : X_* \rightarrow P_s$, and $\Lambda^{\#} : X_* \rightarrow P_s$ such that $\Lambda \mid X_\mu = \Lambda^{\#} \mid X_\mu$ for μ sufficiently large, it follows that $\Lambda = \Lambda^{\#}$ (i.e. $\Lambda(W) = \Lambda^{\#}(W)$). Indeed, let Λ be given by sections $\sigma_0, \cdots, \sigma_s$ of $\boldsymbol{L}_*$ and $\Lambda^{\#}$ by sections $\tau_0, \cdots, \tau_s$ of $\boldsymbol{L}_*{}^{\#}$. Then, since $\boldsymbol{L}_* \mid X_\mu \cong \boldsymbol{L}_*{}^{\#} \mid X_\mu$, there exists a (unique) bundle equivalence $\phi : \boldsymbol{L}_* \rightarrow \boldsymbol{L}_*{}^{\#}$ and furthermore $\phi(\sigma_j) \mid X_\mu = \tau_j \mid X_\mu$. By the uniqueness of the power series expansions, it follows that $\phi(\sigma_j) = \tau_j$ as sections of $\boldsymbol{L}_*{}^{\#}$. Thus $\Lambda = \Lambda^{\#}$.

Let now $\hat{\boldsymbol{L}}$ be the formal line bundle on $\hat{X}$ determined by $\boldsymbol{L}_*$, and set $\hat{\boldsymbol{L}}' = \hat{f}^{-1}(\hat{\boldsymbol{L}})$.

Let μ be sufficiently large: Set $\boldsymbol{L}'_\mu = \hat{\boldsymbol{L}}' \mid X'_\mu$ and let $\mu^{\sigma_0}, \cdots, \mu^{\sigma_s}$ be the sections of $\hat{\boldsymbol{L}}'_\mu$ over X'_μ determined by $\hat{f}$. By Theorem I, we may find a bundle $\boldsymbol{L}'_* \rightarrow X'_*$ and sections $\sigma'_0, \cdots, \sigma'_s$ of L'_* such that $\boldsymbol{L}_*' \mid X_\mu = \boldsymbol{L}'_\mu$ and $\sigma'_j \mid \boldsymbol{L}'_\mu = {}_\mu\sigma_j$. Furthermore, the sections $\sigma'_0, \cdots, \sigma'_s$ give a projective embedding $\Lambda' : X'_* \rightarrow P_s$.

Suppose now that we look at $\boldsymbol{L}'_{\mu+1} = \boldsymbol{L}' \mid X'_{\mu+1}$ and ${}_{\mu+1}\sigma_0, \cdots, {}_{\mu+1}\sigma_s$ as the sections of $\boldsymbol{L}'_{\mu+1}$ determined by $\hat{f}$. By Theorem I, we may find a bundle $\boldsymbol{L}_*{}^{\#} \rightarrow X'_*$ and sections $\tau'_0, \cdots, \tau'_s$ of $\boldsymbol{L}_*{}^{\#}$ such that $\boldsymbol{L}_*{}^{\#} \mid X'_{\mu+1} = \boldsymbol{L}'_{\mu+1}$ and $\tau_j \mid X_{\mu+1} = {}_{\mu+1}\sigma_j$. If then $\Lambda^{\#} : X'_* \rightarrow P_s$ is the corresponding projective embedding, it follows from the above remark that $\Lambda' = \Lambda^{\#}$. Thus, for every μ, $\hat{f} \mid X'_\mu = \Lambda' \mid X'_\mu$ and so $\hat{f} = \Lambda' : W' \rightarrow \Lambda'(W) \subset P_s$. This says precisely that $\hat{f}$ is induced by an actual equivalence of germs of embeddings. Q.E.D.

Proof of (ii). Let $\boldsymbol{E} \rightarrow X$ be a holomorphic vector bundle and suppose that we have an extension $\boldsymbol{E}_\mu \rightarrow X_\mu$, where μ is sufficiently large. Then there exists a formal bundle $\hat{\boldsymbol{E}} \rightarrow \hat{X}$ such that $\hat{\boldsymbol{E}} \mid X_\mu = \boldsymbol{E}_\mu$. We want to prove that $\hat{\boldsymbol{E}}$ converges.

Let $\boldsymbol{L} \rightarrow X$ be the ample line bundle which has an ample extension $\boldsymbol{L}_* \rightarrow X_*$. For N sufficiently large, $\boldsymbol{E} \otimes \boldsymbol{L}^N$ will be ample, and we thus get an embedding $\sigma_\mu : X_\mu \rightarrow G(m,r) \rightarrow P_{m'}$, for some m and m', such that $\sigma_\mu{}^{-1}(\boldsymbol{E}(m,r)) = \boldsymbol{E}_\mu \otimes \boldsymbol{L}_\mu{}^N$. Here $G(m,r)$ is the Grassman variety of m-planes in $\boldsymbol{C}^{m+r}$ and $\boldsymbol{E}(m,r) \rightarrow G(m,r)$ is the universal bundle, with fibre $\boldsymbol{C}^r$, over $G(m,r)$. If we can extend σ_μ to a mapping $\sigma_* : X_* \rightarrow G(m,r)$, then we may set $\boldsymbol{E}_* = \sigma_*{}^{-1}(\boldsymbol{E}(m,r)) \otimes \boldsymbol{L}_*{}^{-N}$ and $\boldsymbol{E}_*$ is an extension to X_* of $\boldsymbol{E}_\mu \rightarrow X_\mu$.

By the same reasoning as above, we may extend σ_μ to a projective embedding $\tilde{\sigma}_* : X_* \rightarrow P_{m'}$. On the other hand, $\tilde{\sigma}_*(X)_\mu = \sigma_\mu(X) \subset G(m,r) \subset P_{m'}$ and the normal bundle $\boldsymbol{N}$ of X in X_* is a sub-bundle or the normal bundle of X in $G(m,r)$. We then consider the complete continuous system

$\{X_t\}$ of X in $G(m,r)$ corresponding to the sections of $\boldsymbol{N}$. The family $\{X_t\}$ gives a local sub-manifold $W' \subset G(m,r)$ and $X \subset W'$.* Furthermore, since $\tilde{\sigma}_* \mid X_\mu = \sigma_\mu$, their germs of embeddings $X \subset W'$ and $X \subset W$ are equivalent up to order μ. Thus, if μ is sufficiently large, they are equivalent by a mapping $g: W \rightarrow W'$. If we set $\sigma_* = g \cdot \tilde{\sigma}_*$, then $\sigma_*: W \rightarrow G(m,r)$ is our required mapping.

Proof of (iii). Suppose that $\boldsymbol{E}_* \rightarrow X_*$ and $\tilde{\boldsymbol{E}}_* \rightarrow X_*$ are vector bundles such that $\boldsymbol{E}_* \mid X_\mu \cong \tilde{\boldsymbol{E}}_* \mid X_\mu$ for μ large. This equivalence is given by an element $\gamma_\mu \in H^0(X, \Omega(\mathrm{Hom}(\boldsymbol{E}_* \mid \tilde{X}_\mu, \tilde{\boldsymbol{E}}_* \mid X_\mu)))$. By Theorem I we may find an extension $\gamma_* \in H^0(X, \Omega(\mathrm{Hom}(\boldsymbol{E}_*, \tilde{\boldsymbol{E}}_*)))$ of γ_μ, and clearly γ_* gives a bundle equivalence between $\boldsymbol{E}_* \rightarrow X_*$ and $\tilde{\boldsymbol{E}}_* \rightarrow X_*$. Q.E.D.

IV. Some Examples and Remarks about Positive Embeddings.

1. Proof of Theorem I for positively embedded hypersurfaces. Let $X \subset W$ be a positive embedding of codimension 1. By Propositions 3.5 and 3.4 (neither of which use Theorem I), there exists a projective embedding $\Phi: W \rightarrow P_N$ given by $|\boldsymbol{L}^{\mu_0}|$ $(\boldsymbol{L} = [X])$ for μ_0 large enough. On the other hand, by VIII.2 and VIII.3, there exists $W^\# \subset W$ and a real non-negative function r defined on W such that $\partial\bar{\partial} r$ has, in $W - W^\#$, n negative eigenvalues. Here $n = \dim X$ and we assume that $n \geqq 2$. Thus $\dim H^0(W, \Omega(\boldsymbol{E})) < \infty$ for any line bundle $\boldsymbol{E} \rightarrow W$. Then, by the argument of Siegel (cf. [1] for the details) we find that the field $\mathfrak{M}$ of meromorphic functions on W is an algebraic function field of transcendence degree less than or equal to $n+1 = \dim W$. Let $V \subset P_N$ be the smallest algebraic variety containing $\Phi(W)$ and $\mathfrak{K}$ the field of rational functions on V. Then the restriction mapping $\mathfrak{K} \rightarrow \mathfrak{M}$ is obviously an injection and so $\dim V = \deg \mathfrak{K} \leqq \deg \mathfrak{M} \leqq \dim W$; i.e. $\dim V = \dim W$. Thus V is an algebraic compactification of $\Phi(W) \subset P_N$. Furthermore, by construction, $\mu_0 \Phi(X) \subset V$ is a hyperplane section.

Let $\boldsymbol{L} \rightarrow V$ be the line bundle $[\Phi(X)]$. For $\boldsymbol{E} \rightarrow V$ any holomorphic bundle, set $\boldsymbol{E}(\mu) = \boldsymbol{E} \otimes \boldsymbol{L}^{*\mu}$. By Theorems A and B for algebraic varieties, we get

$$H^q(V, \Omega(\boldsymbol{E}(\mu))) = 0 \text{ for } 0 \leqq q < n+1, \mu \geqq \mu_0(\boldsymbol{E}). \tag{4.1}$$

Now set $Z = V - \Phi(W')$ where $W^\# \subset W' \subset W$. If Ψ is the family of all closed sets in Z, we have an exact cohomology sequence

---

* In fact, W' is the image of a complex manifold V' under a holomorphic mapping $F': V' \rightarrow G(m,r)$ such that F'_* has everywhere constant rank (cf. Thtorem 3.1).

$$H_{\Psi}{}^{q}(Z,\Omega(\boldsymbol{E}(\mu)))\to H^{q}(V,\Omega(\boldsymbol{E}(\mu)))\to H^{q}(W',\Omega(\boldsymbol{E}(\mu)))\to H_{\Psi}{}^{q+1}(Z,\Omega(\boldsymbol{E}(\mu)))\to\cdots. \tag{4.2}$$

By [3], §20, $H_{\Psi}{}^{q}(Z,\Omega(\boldsymbol{E}(\mu)))=0$ for $0\leqq q<n+1$ and then (4.1) and (4.2) give

$$H^{q}(W',\Omega(\boldsymbol{E}(\mu)))=0 \text{ for } 0\leqq q<n, \mu\geqq\mu_0(\boldsymbol{E}). \tag{4.3}$$

Now (4.3) gives Theorem I, at least for those bundles $\boldsymbol{E}\to W$ which extend to V.*

COROLLARY. *Theorem II holds for positive embeddings of codimension* 1.

Proof. We must show that line bundles extend from $\Phi(W)$ to V. But this follows from Proposition (1.5).

2. Rigidity of some special embeddings. Let $\phi: X\to W$ and ϕ': $X\to W'$ be positive embeddings. Assuming Theorem I and that $\dim X\geqq 3$, we may find projective embeddings Φ: $W\to P_N$, Φ': $W'\to P_{N'}$. Now the same argument as in §4.1 above shows that there exist algebraic compactifications $V\subset P_N$, $V'\subset P_{N'}$ of $\Phi(W)$, $\Phi'(W')$ respectively. Thus $\Phi(W)\subset V$, $\dim\Phi(W)=\dim V$, and the restriction of rational functions from V to $\Phi(W)$ is an isomorphism of function fields. Consequently, any equivalence of germs F: $W\to W'$ will induce a meromorphic mapping Λ: $V\to V'$. Now it sometimes happens that Λ will automatically be bi-holomorphic (e.g. if $V=V'$ and V is a minimal variety such as P_n, an abelian variety, or an algebraic manifold uniformized by a bounded domain.* In this situation one may discuss, by direct geometric methods the rigidity problem. We give two examples of this:

(i) Projective embeddings. Let $P=P_n$ and let ϕ: $X\to P$, ϕ': $X\to P$ be two positive embeddings.

PROPOSITION 4.1. *Suppose that we have an equivalence between* 1-*neighborhoods* h: $X'_1\to X_1$ *such that* $h\mid X$ *is homotopic to the identity. Then there exists a projective transformation* $g\in SL(N+1,\boldsymbol{C})$ *such that* $g\circ\phi'=\phi$ *and* $g\mid X'_1=h$.

Proof. Let $\boldsymbol{H}\to P$ be the hyperplane bundle. We recall the exact sequence

$$0\to\boldsymbol{l}\to(\boldsymbol{H})^{N+1}\to\boldsymbol{T}(P)\to 0 \tag{4.4}$$

---

* We remark that $H^{n+1}(W',\Omega(\boldsymbol{F}))=0$ for any holomorphic bundle $\boldsymbol{F}\to W'$.

* The point here is that every automorphism of the function field induces one of the variety.

where $\boldsymbol{l}$ is the trivial bundle and $(\boldsymbol{H})^{N+1} = \underbrace{\boldsymbol{H} \oplus \cdots \oplus \boldsymbol{H}}_{N+1}$.

Now $\Omega(\boldsymbol{T}(P)^*) = \Omega^1(P) =$ the germs of holomorphic 1-forms on P, and (4.4) is uniquely given by $\boldsymbol{l}$, $\boldsymbol{T}(P)$, and the extension elsment $\omega \in H^1(P, \Omega^1(P))$. Clearly ω is the fundamental form on P.

Let $\boldsymbol{T} = \boldsymbol{T}(P) \mid \phi(X)$, $\boldsymbol{T}' = \boldsymbol{T}(P) \mid \phi'(X)$ and denote by $\boldsymbol{N}$, $\boldsymbol{N}'$ the respective normal bundles of $\phi(X)$, $\phi'(X)$ in P. Then $h\colon X'_1 \to X_1$ is given by a bundle mapping $h\colon \boldsymbol{T}' \to \boldsymbol{T}$ which takes

$$0 \to \boldsymbol{T}(X) \to \boldsymbol{T}' \to \boldsymbol{N}' \to 0 \tag{4.5$'$}$$

onto

$$0 \to \boldsymbol{T}(X) \to \boldsymbol{T} \to \boldsymbol{N} \to 0. \tag{4.5}$$

Write f for the automorphism of X which h induces.

Letting $\boldsymbol{L} = \boldsymbol{H} \mid \phi(X), \boldsymbol{L}' = \boldsymbol{H} \mid \phi'(X)$, we want to show that h extends to a mapping of

(4.6)′

$$\begin{tikzcd}
 & & 0 \arrow[d] & & \\
 & & \boldsymbol{l} \arrow[d] & & \\
 & & (\boldsymbol{L}')^{N+1} \arrow[d] & & \\
0 \arrow[r] & \boldsymbol{T}(X) \arrow[r] & \boldsymbol{T}' \arrow[r] \arrow[d] & \boldsymbol{N}' \arrow[r] & 0 \\
 & & 0 & &
\end{tikzcd}$$

onto

(4.6)

$$\begin{tikzcd}
 & & 0 \arrow[d] & & \\
 & & \boldsymbol{l} \arrow[d] & & \\
 & & (\boldsymbol{L})^{N+1} \arrow[d] & & \\
0 \arrow[r] & \boldsymbol{T}(X) \arrow[r] & \boldsymbol{T} \arrow[r] \arrow[d] & \boldsymbol{N} \arrow[r] & 0 \\
 & & 0. & &
\end{tikzcd}$$

Now (4.6)′ is given by $\gamma' \in H^1(X, \Omega(\boldsymbol{T}'^*))$, (4.6) is given by $\gamma \in H^1(X, \Omega(\boldsymbol{T}^*))$, and h will extend to a mapping of (4.6)′ onto (4.6) if $h^*(\gamma) = \gamma'$ where $h^*\colon H^1(X, \Omega(\boldsymbol{T}^*)) \to H^1(X, \Omega(\boldsymbol{T}'^*))$ is the cohomology map.

Since $\boldsymbol{N}$ and $\boldsymbol{N}'$ are positive and $\dim X \geqq 2$,

$$H^1(X, \Omega(\boldsymbol{N}^*)) = 0 = H^1(X, \Omega(\boldsymbol{N}'^*)).$$

Letting $\Omega^1 = \Omega(\boldsymbol{T}(X)^*)$, we have

$$\begin{tikzcd}
& 0 \arrow[r] & H^1(X,\Omega(\boldsymbol{T}'^*)) \arrow[r, "\sigma"] & H^1(X,\Omega^1) \\
H^1(P,\Omega^1(P)) \arrow[urr] \arrow[urrr, "\alpha"] \arrow[drr] \arrow[drrr, "\beta"] & & & \\
& 0 \arrow[r] & H^1(X,\Omega(\boldsymbol{T}^*)) \arrow[uu, "h^*"] \arrow[r, "\tau"] & H^1(X,\Omega^1) \arrow[uu, "f^*"']
\end{tikzcd}$$

Because $\sigma(\gamma') = \alpha(\omega)$ and $\tau(\gamma) = \beta(\omega)$, to prove that $h^*(\gamma) = \gamma'$ it will suffice to show that $f^*\beta(\omega) = \alpha(\omega)$. But this is true since $f\beta \sim \alpha$ and X is Kähler. Thus h induces a bundle mapping $h: \boldsymbol{L}' \to \boldsymbol{L}$ such that $h((4.6)') = (4.6)$.

The embedding $\phi: X \to P$ is given by a linear subspace $S \subset H^0(X,\Omega(\boldsymbol{L}))$; ϕ' is given by $S' \subset H^0(X,\Omega(\boldsymbol{L}'))$. To prove Proposition 4.1, we must show that $h_*(S') = S$ where $h_*: H^0(\Omega(\boldsymbol{L}')) \to H^0(\Omega(\boldsymbol{L}))$ is the cohomology map. Let $n+1 = \dim H^0(\Omega(\boldsymbol{L}))$. There are inclusions $P \subset |\boldsymbol{L}| \cong P_n$, $P \subset |\boldsymbol{L}'| \cong P'_n$. Furthermore, the following diagram commutes:

$$\begin{tikzcd}
& {|S|} \arrow[r] & P_n \\
X \arrow[ur, "\phi"] \arrow[dr, "\phi'"'] & & \\
& {|S'|} \arrow[r] & P'_n \arrow[uu, "h_*"']
\end{tikzcd}$$

Finally, $h_*(|S'|) \subset P_n$ is a linear subspace which osculates along $\phi(X)$ to $|S|$ with at least first order contact. Thus $h_*(|S'|) = |S|$. Q. E. D.

Remark. As promised, we give now a proof of Lemma 3.1. We have the exact bundle sequence

$$0 \to \boldsymbol{T}(P_n)(\mu) \to \boldsymbol{T}(\mu) \to \boldsymbol{H}^{*\mu-1} = 0 \qquad (\mu \geqq 1)$$

and, since $H^1(P_n, \Omega(\boldsymbol{H}^{*\mu-1})) = 0$ $(\mu \geqq 1)$, it will suffice to show

$$H^1(P_n, \Omega(\boldsymbol{T}(P_n) \otimes \boldsymbol{H}^{*\mu})) = 0$$

for $\mu \geqq 1$. From (4.4) we have

$$0 \to \boldsymbol{H}^{*\mu} \to (\boldsymbol{H}^{*\mu-1})^{n+1} \to \boldsymbol{T}(P_n)(\mu) \to 0. \tag{4.7}$$

From the exact cohomology sequence of (4.7) it follows that

$$H^1(X, \Omega(\boldsymbol{T}(P_n)(\mu))) = 0 \qquad (\mu > 0),$$

provided that $n > 2$.

Q. E. D.

(ii) Embeddings in abelian varieties. Let X be a compact algebraic manifold. If we fix a point $x_0 \in X$, then there is defined an Abelian variety A_X, the *Albanese variety* of X, and regular mapping $\psi \colon X \to A_X$ which has the following universal property: Given any complex torus T and a mapping $\xi \colon X \to T$, there exists $t_\xi \in T$ and a homomorphism $\theta \colon A_X \to T$ such that $\xi(x) = \theta\psi(x) + t_\xi$ $(x \in X)$. We say that ξ is normalized if $\xi(x_0)$ is the identity in T.

LEMMA 4.1. *Let A be an abelian variety and $\phi \colon X \to A$ a positive normalized embedding. Then A_X is a finite covering torus of A and θ is the projection homomorphism.*

Proof. Letting $\boldsymbol{T} = \boldsymbol{T}(A) \mid \phi(X)$ and $\Omega^1 = \Omega(\boldsymbol{T}(X)^*)$, we get from $0 \to \boldsymbol{T}(X) \to \boldsymbol{T} \to \boldsymbol{N} \to 0$ the sequence

$$0 \to H^0(X, \Omega(\boldsymbol{T}^*)) \xrightarrow{\gamma} H^0(X, \Omega^1) \to 0. \tag{4.8}$$

Now $\boldsymbol{T}_e(A_X)^* \cong H^0(X, \Omega^1)$, $\boldsymbol{T}_e(A)^* \cong H^0(A, \Omega^1(A)) \cong H^0(X, \Omega(\boldsymbol{T}^*))$, and $\gamma \colon \boldsymbol{T}_e(A)^* \to \boldsymbol{T}_e(A_X)^*$ is just the transpose of θ_*, the differential of θ at the identity. The Lemma then follows from (4.8). Q.E.D.

From Lemma 4.1 we may easily prove

PROPOSITION 4.2. *Let A be an abelian variety. Then a positive embedding $\phi \colon X \to A$ $(\dim X \geqq 2)$ is uniquely determined, up to an automorphism of A, by its normal bundle.*

3. On ample hypersurfaces. We discuss now some applications of the rigidity theorem; the main one of these is

THEOREM 4.1. *Let X be an algebraic manifold of dimension $n \geqq 2$. Then there exists an integer $\mu_0 = \mu_0(X)$ such that the following is true: If W, W' are any two nonsingular varieties for which we have embeddings $\phi \colon X \to W$, $\phi' \colon X \to W'$ as hyperplane section, and if we have an equivalence $X_\mu \cong X'_\mu$ for $\mu \geqq \mu_0$, then there exists a biregular mapping $\Phi \colon W \to W'$ such that*

$$\begin{tikzcd}
 & W \arrow[dd, "\Phi"] \\
X \arrow[ur, "\phi"] \arrow[dr, "\phi'"'] & \\
 & W'
\end{tikzcd}$$

commutes.

Remark. Briefly we may say that an algebraic variety of dimension $\geqq 3$ is determined by a finite neighborhood of a hyperplane section.

Let X be an algebraic variety of dimension $n \geqq 2$ and W a non-singular variety of dimension $n+1$. We say that an embedding $\phi: X \to W$ is *ample* if, and only if, the line bundle $\boldsymbol{L}_\phi = [\phi(X)]$ on W is positive. Equivalently, $|\boldsymbol{L}_\phi{}^\mu|$ for large μ should give a projective embedding. Clearly, if ϕ is ample and if $C \subset W$ is any curve, then $\phi(X) \cap C \neq \emptyset$. Also the normal bundle $\boldsymbol{N}_\phi$ to $\phi(X) \subset W$ is positive. Conversely, we have

PROPOSITION 4.3. *Let $\phi: X \to W$ be an embedding such that $\boldsymbol{N}_\phi$ is positive and such that $\phi(X) \cap C \neq \emptyset$ for any curve $C \subset W$. Then ϕ is ample.*

Proof. By the same argument as in IV.1, there exists a neighborhood U of $\phi(X) \subset W$ and a μ_0 such that:

(i) For $\mu \geqq \mu_0$, we have
$$H^0(U, \Omega(\boldsymbol{L}_\phi{}^\mu)) \to H^0(X, \Omega(\boldsymbol{N}_\phi{}^\mu)) \to 0;$$

(ii) $|\boldsymbol{L}_\phi{}^{\mu_0}|$ gives a projective embedding $\Phi: U \to P_N$.

By VIII.3 below, we may further assume:

(iii) The E. E. Levi form in a neighborhood of ∂U has everywhere n positive eigenvalues. By (iii) we see that, for any holomorphic bundle $\boldsymbol{E} \to W$, the restriction mapping gives an isomorphism

$$0 \to H^q(W, \Omega(\boldsymbol{E})) \to H^q(U, \Omega(\boldsymbol{E})) \to 0$$

for $q < n-1$. Since $n \geqq 2$, we get, in particular,

$$0 \to H^0(W, \Omega(\boldsymbol{E})) \to H^0(U, \Omega(\boldsymbol{E})) \to 0.$$

Thus we can extend Φ to a holomorphic mapping $\Phi: W \to P_N$.

Consider now the differential Φ_* of $\Phi: W \to P_N$. Since $\Phi(X) \cap C \neq \emptyset$ for any curve C, it follows that Φ_* is singular at finitely many points. But then it follows easily that $\boldsymbol{L}_\phi \to W$ is positive. Q. E. D.

COROLLARY 4.1. *The following are equivalent:*

(i) *$\phi: X \to W$ is ample;*

(ii) *$\boldsymbol{N}_\phi \to X$ is positive and $\phi(X) \cdot C > 0$ for every irreducible curve C.*

(iii) *There exists a positive bundle $\boldsymbol{L} \to W$ and $\sigma \in H^0(W, \Omega(\boldsymbol{L}))$ such that $\phi(X) = \{w \in W : \sigma(w) = 0\}$.*

Remark. Proposition 4.3 is closely related to a theorem of Nakai [16] which reads:

(*) Let $X \subset W$ be a hypersurface such that $X^r \cdot Y > 0$ for any subvariety $Y \subset W$ of dimension $r < n+1$. Then X is ample.

In order to prove (*) from the Corollary above, we introduce the seemingly more general statement

(#) Let $\boldsymbol{E} \to W$ be a line bundle such that $(\boldsymbol{E})^r \cdot Y > 0$ for any r-dimensional sub-variety $Y \subset W$. Then $\boldsymbol{E}$ is positive.

Remark. $(\boldsymbol{E})^r \cdot Y$ is defined in the obvious way using the Chern class of $\boldsymbol{E}$.

In order to prove (*), using the Corollary above it will suffice to show that the normal bundle $\boldsymbol{N} \to X$ is positive. But this follows from (#) in one dimension lower, since for $Y \subset X$ an r-dimensional variety, $(\boldsymbol{N})^r \cdot Y = X^r \cdot Y > 0$ because of $\boldsymbol{N} = [X] \mid X$. On the other hand, (*) is true for $n = 1$ since $\deg \boldsymbol{N} = X \cdot X > 0$. Thus, by induction, we may assume that (*) holds in dimension n, (#) holds in dimension $n-1$, and show that (#) holds in dimension n. The argument for this is not difficult and will be omitted.

Proof of Theorem 4.1. Assuming $X_\mu \cong X'_\mu$ for $\mu \geqq \mu_0$, the integer for the rigidity theorem, we find neighborhoods U, U' of $\phi(X)$, $\phi'(X)$ in W, W' and a biholomorphic map $\Phi \colon U \to U'$ such that

$$\begin{tikzcd} & U \arrow[dd, "\Phi"] \\ X \arrow[ur, "\phi"] \arrow[dr, "\phi'"'] & \\ & U' \end{tikzcd}$$

commutes. Furthermore, we may assume that U and U' are embedded in a projective spaces of the same dimension by $|\boldsymbol{L}_{\phi}{}^{\mu_0}|$ and $|\boldsymbol{L}_{\phi'}{}^{\mu_0}|$ respectively. But then Φ extends to a meromorphic mapping $\Phi \colon W \to W'$ and the argument of Proposition 4.3 shows that Φ is everywhere defined. The same holds for Φ^{-1} and thus $\Phi \colon W \to W'$ is biregular. Q. E. D.

4. A partial generalization to higher codimension. Let $G(m,r)$ be the Grassmann variety of m-planes in $\boldsymbol{C}^{m+r}$. Over $G(m,r)$ there is the universal bundle $\boldsymbol{F}(m,r)$ whose fibre over $\xi \in G(m,r)$ is just the m-dimensional vector space ξ. There is the usual exact sequence

$$0 \to \boldsymbol{F}(m,r) \to \boldsymbol{C}^{m+r} \to \boldsymbol{E}(m,r) \to 0 \tag{4.9}$$

where $\boldsymbol{C}^{m+r}$ is the trivial bundle $G(m,r) \times \boldsymbol{C}^{m+r}$. We remark that $\dim G(m,r) = mr$ and that

$$\boldsymbol{T}(G(m,r)) \cong \operatorname{Hom}(\boldsymbol{F}(m,r),\boldsymbol{E}(m,r)). \tag{4.10}$$

One main difficulty in generalizing IV. 3 above is the following:

PROPOSITION 4. 4 [8]. *$\boldsymbol{E}(m,r) \to G(m,r)$ is weakly positive if, and only if, $r=1$.*

At the moment, we shall thus settle for a generalization of the implication (iii) $\Rightarrow$ (ii) in Corollary 4. 1. Let then $\boldsymbol{E} \to W$ be a holomorphic bundle with fibre $\boldsymbol{C}^r$ over the compact variety W of dimension $n+r$ $(r \geqq 1)$. Let further $\dim H^0(W,\Omega(\boldsymbol{E})) = m+r$ $(m \geqq 0)$.

Definition. We shall say that $\boldsymbol{E} \to W$ is *ample* if: (i) the global sections in $H^0(W,\Omega(\boldsymbol{E}))$ generate every fibre $\boldsymbol{E}_w$; and (ii) if $w_0 \in W$ and if $s_1,\cdots,s_m \in H^0(W,\Omega(\boldsymbol{E}))$ are a basis for the sections which vanish at w_0, then the differentials $(ds_a)(w_0)$ $(1 \leqq a \leqq m)$ should span $\boldsymbol{E}_{w_0} \otimes \boldsymbol{T}_{w_0}(W)^*$.

Remark. If $\boldsymbol{E} \to W$ is ample, then it is weakly positive (VI. 5). If, conversely, $\boldsymbol{E} \to W$ is weakly positive, then $\boldsymbol{E}^\mu \to W$ is ample for $\mu \geqq \mu_0$.

By definition, if $\boldsymbol{E} \to W$ is ample, then $\dim H^0(W,\Omega(\boldsymbol{E})) \geqq r + r(\dim W)$. Thus $\boldsymbol{E}(m,r)$ is not ample if $r>1$. On the other hand, $\boldsymbol{T}(P_n)$ is ample (since the projective group acts transitively on the tangent directions to P_n); observe that $\dim H^0({}_n,\boldsymbol{T}(P_n)) = n^2+2n \geqq n^2+n$.

If $\boldsymbol{E} \to W$ is ample, we have the exact sequences

$$0 \to \boldsymbol{F} \to W \times H^0(W,\Omega(\boldsymbol{E})) \xrightarrow{\rho} \boldsymbol{E} \to 0 \tag{4.11}$$

$$\boldsymbol{F} \xrightarrow{\sigma} \boldsymbol{E} \otimes \boldsymbol{T}(W)^* \to 0 \tag{4.12}$$

where $\rho(w,s) = s(w) \in \boldsymbol{E}_w$ for $w \in W$, $s \in H^0(W,\Omega(\boldsymbol{E}))$ and σ is given by $\sigma_w(s) = (ds)(w)$ since $\boldsymbol{F}_w \cong$ sections $s \in H^0(W,\Omega(\boldsymbol{E}))$ with $s(w)=0$.

Now the mapping $w \to \boldsymbol{F}_w \subset H^0(W,\Omega(\boldsymbol{E}))$ defines $\phi\colon W \to G(m,r)$ such that $\phi^{-1}(\boldsymbol{E}(m,r)) = \boldsymbol{E}$ and $\phi^{-1}(\boldsymbol{F}(m,r)) = \boldsymbol{F}$. The differential ϕ_* of ϕ is a global section of

$$\operatorname{Hom}(\boldsymbol{T}(W),\phi^{-1}(\boldsymbol{T}(G(m,r)))) = \operatorname{Hom}(\boldsymbol{T}(W),\operatorname{Hom}(\boldsymbol{F},\boldsymbol{E}))$$

by (4. 10). Thus $\phi^* \in \operatorname{Hom}(\operatorname{Hom}(\boldsymbol{E},\boldsymbol{F}),\boldsymbol{T}(W)^*)$. On the other hand, there is a natural isomorphism

$$\operatorname{Hom}(\operatorname{Hom}(\boldsymbol{E},\boldsymbol{F}),\boldsymbol{T}(W)^*) \cong \operatorname{Hom}(\boldsymbol{F},\boldsymbol{E} \otimes \boldsymbol{T}(W)^*)$$

and it is easy to see that, under this isomorphism, ϕ^* corresponds to σ given in (4. 12). That is, for all $\lambda \in \boldsymbol{T}(W)$, $e^* \in \boldsymbol{E}^*$, $f \in \boldsymbol{F}$, we have

$$\langle \lambda, \phi^*(e^* \otimes f)\rangle = \langle e^* \otimes \lambda, \sigma(f)\rangle.$$

In particular, if σ is onto, then ϕ^* is onto but the converse is *not* true unless $r=1$. Thus $\phi\colon W\to G(m,r)$ is an embedding but this does *not* make $\phi^{-1}(\boldsymbol{E}(m,r))$ ample unless $r=1$. In fact, we can prove

PROPOSITION 4.5. *$\phi^{-1}(\boldsymbol{E}(m,r))$ is ample if $\phi\colon W\to G(m,r)$ is an embedding and if the holomorphic second fundamental form of $\phi(W)\subset G(m,r)$* (cf. VI.3) *is non-degenerate.*

Thus, in order that $\phi^{-1}(\boldsymbol{E}(m,r))$ *be ample,* $\phi(W)$ must be sufficiently twisted in $G(m,r)$.

THEOREM 4.2. *Let $s\in H^0(W,\Omega(\boldsymbol{E}))$ and set $X=\{w\in W\colon s(w)=0\}$. Then, assuming $\boldsymbol{E}\to W$ is ample, $X\cdot Y>0$ for every irreducible variety $Y\subset W$ of dimension r.*

Remarks. (i) The normal bundle N of $X\subset W$ is given by $\boldsymbol{N}=E\mid X$. Thus $\boldsymbol{N}\to X$ has whatever positivity is present in $\boldsymbol{E}\to W$. In particular, $\boldsymbol{N}$ is weakly positive.

(ii) For $r=1$, this includes the implications (iii) $\Rightarrow$ (ii) of Corollary 4.1. Indeed, $\boldsymbol{L}^{\mu}\to X$ is ample for large μ, and $\phi(X)\cdot C=1/\mu(\boldsymbol{L}^{\mu})\cdot C>0$.

(iii) Theorem 4.2 does *not* follow from the fact that $w\to\boldsymbol{F}_w$ gives an embedding $\phi\colon W\to G(m,r)$. In fact, Theorem 4.2 is false for $\boldsymbol{E}(m,r)\to G(m,r)$.

(*Proof.* For $a=(a_1,\cdots,a_r)$ with $0\leqq a_1<\cdots<a_r\leqq m+r$, the (generic) Schubert variety $\Omega_a\subset G(m,r)$ is defined ([10], § XIV.2); $\dim\Omega_a=\sum_{i=1}^{r}a_i-\frac{1}{2}r(r-1)$. If $b=(b_1,\cdots,b_r)$ is such that $\dim\Omega_a+\dim\Omega_b=mr$, then $\Omega_a\cdot\Omega_b\geqq 0$ and $\Omega_a\cdot\Omega_b>0$ if, and only if, $a_i+b_{r-i}\geqq m+r$ $(1\leqq i\leqq r)$. For $0<s\leqq r$, we let Ω_s be the Schubert variety with the symbol $(\overbrace{m,\cdots,m+s-1}^{s},\ \overbrace{m+s+1,\cdots,m+r}^{r-s})$. Then $\dim\Omega_s=mr-s$, and in fact Ω_s is dual to $c_s(\boldsymbol{E})$. Now we have $\Omega_r\cdot\Omega_b=\int_{\Omega_b}c_r(\boldsymbol{E})=0$ if say $b=(0,2,3,\cdots,r-1,r+1)$ (for the notations see the proof of Theorem 4.2). At the last step we have, of course, used that $r>1$.)

Proof of Theorem 4.2. Let $c(\boldsymbol{E})=c_0+\cdots+c_r$ $(c_i\in H^{2i}(W,\boldsymbol{Z}))$ be the total Chern class of $\boldsymbol{E}\to W$. Then, if $y\in H_{2r}(W,\boldsymbol{Z})$ is the homology class of Y,

$$X\cdot Y=\langle c_r,y\rangle \tag{4.13}$$

where $\langle\ ,\ \rangle \colon H^{2r}(W, \boldsymbol{Z}) \otimes H_{2r}(W, \boldsymbol{Z}) \to \boldsymbol{Z}$ is the usual pairing. In order to prove Theorem 4.1, we shall prove the more general

THEOREM 4.2. *Let $Y \subset W$ be a irreducible variety of dimension q, $0 < q \leqq r$. Then, with the above notation,*

$$\langle c_q, y \rangle > 0. \tag{4.14}$$

Now E fits in the exact bundle sequence

$$0 \to \boldsymbol{F} \to \boldsymbol{C}^{m+r} \xrightarrow{\rho} \boldsymbol{E} \to 0 \tag{4.15}$$

where $\boldsymbol{C}^{m+r}$ is the trivial bundle $W \times H^0(W, \Omega(\boldsymbol{E}))$. The flat metric in the middle term induces then a metric in $\boldsymbol{E}$ and there are then defined real (q, q) forms θ_q on W with $d\theta_q = 0$ and such that θ_q represents the de Rham class of c_q. Then

$$\langle c_q, y \rangle = \int_Y \theta_q \tag{4.16}$$

and it will suffice to show that $\theta_q \mid Y$ gives a volume element. This is a local differential-geometric question, and we shall use the concepts and notations developed in § VI below.

The curvature form Θ in $\boldsymbol{E}$ is locally written as $\Theta = \{\Theta_{\rho\sigma}\}$ where

$$\Theta_{\rho\sigma} = \sum \Theta_{\rho\sigma ij} \omega^i \wedge \bar{\omega}^j \qquad (1 \leqq \rho, \sigma, \tau \leqq r;\ 1 \leqq i, j \leqq n + r).$$

Also, the θ_q are uniquely defined by (cf. [5]):

$$\sum_{t=0}^{r} \theta_q t^{r-q} = \det\left(t\boldsymbol{I}^r + \frac{\sqrt{-1}}{2\pi}\Theta\right) \tag{4.17}$$

where $\boldsymbol{I}^r$ is the $r \times r$ unit matrix. Thus, if $S(q) =$ permutation group on q letters,

$$\theta_q = \left(\frac{\sqrt{-1}}{2\pi}\right)^q \sum_{1 \leqq \rho_1 < \cdots < \rho_q \leqq r} \ \sum_{\pi \in S(q)} (-1)^\pi \Theta_{\rho_1 \pi(\rho_1)} \wedge \cdots \wedge \Theta_{\rho_q \pi(\rho_q)} \tag{4.18}$$

Now, by VII.3 below, Θ is given as $\Theta = {}^t b \wedge \bar{b}$ where $b = \{b_{a\rho}\}$ with

$$b_{a\rho} = \sum_i b_{a\rho i} \omega^i \qquad (1 \leqq a \leqq m)$$

is a global holomorphic section of $\operatorname{Hom}(\boldsymbol{E}, \boldsymbol{T}(W)^* \otimes \boldsymbol{F})$ called the second fundamental form of $\boldsymbol{F}$ in $\boldsymbol{C}^{m+r}$ (cf. VI.3 where this terminology is justified). In the present case, b is just ϕ^* where $\phi \colon W \to G(m, r)$ is the embedding of W in the Grassmannian such that $\phi^{-1}(4.9) = (4.15)$. Explicitly,

$$\Theta_{\rho\sigma ij} = \sum_{a=1}^{m} b_{a\rho i}\bar{b}_{a\sigma j}, \tag{4.19}$$

Given any point $w_0 \in W$, we may find local holomorphic coordinates $z = (z^1, \cdots, z^{n+r})$ centered at w_0, a trivialization of $\boldsymbol{E}$ near w_0, and generators $s_1, \cdots, s_m$ for $\boldsymbol{F}_{w_0} \subset H^0(W, \Omega(\boldsymbol{E}))$ such that

$$b_{a\rho i}(w_0) = \frac{\partial s_a{}^{\rho}}{\partial z^i}\,(w_0). \tag{4.20}$$

Furthermore, if $Y \subset W$ is a non-singular subvariety of dimension q and passing through w_0, we may assume that Y is locally defined by $z^j = 0$ $(q < j \leqq n + r)$ and, since $\boldsymbol{E} \to W$ is ample, we may choose our frames so that

$$b_{1\alpha\beta}(w_0) = \delta_\beta{}^\alpha \qquad (1 \leqq \alpha, \beta \leqq q). \tag{4.21}$$

LEMMA 4.3.

$$\theta_q = \frac{(-1)^{\frac{q(q-1)}{2}}}{q!}\left(\frac{\sqrt{-1}}{2\pi}\right)^q \sum_{\substack{\rho=(\rho_1,\cdots,\rho_q)\\ a=(a_1,\cdots,a_q)}} B_{\rho a} \wedge \bar{B}_{\rho a} \tag{4.22}$$

where $B_{\rho a} = \sum_{\pi \in S(q)} b_{\pi(a_1)\rho_1} \wedge \cdots \wedge b_{\pi(a_q)\rho_q}$ $(1 \leqq \rho_1 < \cdots < \rho_q \leqq r)$ and where $a = (a_1, \cdots, a_q)$ runs over *all* q-tuples of integers selected from $(1, \cdots, m)$.

Proof. $\sum_{\pi \in S(q)} (-1)^\pi \Theta_{\rho_1 \pi(\rho_1)} \wedge \cdots \wedge \Theta_{\rho_q \pi(\rho_q)} =$ (by (4.19))

$$\sum_{\substack{a=(a_1,\cdots,a_q)\\ \pi \in S(q)}} (-1)^\pi b_{a_1\rho_1} \wedge \bar{b}_{a_1\pi(\rho_1)} \wedge \cdots \wedge b_{a_q\rho_q} \wedge \bar{b}_{a_q\pi(\rho_q)}$$

$$= (-1)^{\frac{q(q-1)}{2}} \sum_{a,\pi} (-1)^\pi b_{a_1\rho_1} \wedge \cdots \wedge b_{a_q\rho_q} \wedge \bar{b}_{a_1\pi(\rho_1)} \wedge \cdots \wedge \bar{b}_{a_q\pi(\rho_q)}$$

$$= (-1)^{\frac{q(q-1)}{2}} \sum_{a,\pi} b_{a_1\rho_1} \wedge \cdots \wedge b_{a_q\rho_q} \wedge \bar{b}_{\pi^{-1}a_1\rho_1} \wedge \cdots \wedge \bar{b}_{\pi^{-1}(a_q)\rho_q}$$

$= \dfrac{(-1)^{\frac{q(q-1)}{2}}}{q!} \sum_a B_{\rho a} \wedge \bar{B}_{\rho a}$. Combining this with (4.18), we get

$$\theta_q = \frac{(-1)^{\frac{q(q-1)}{2}}}{q!}\left(\frac{\sqrt{-1}}{2\pi}\right)^q \sum_{\rho,a} B_{\rho a} \wedge \bar{B}_{\rho a},$$

which proves the Lemma.

We may now complete the proof of Theorem 4.2′. Let Y pass through w_0 and choose local frames $\omega^1, \cdots, \omega^{n+r}$ such that $\omega^j = 0$ on Y for $q < j$

$\leqq n+r$. Furthermore, choose a local frame for $\boldsymbol{E}$ so that (4.21) is satisfied. Now let $a_0=(1,\cdots,1)$ and $\rho_0=(1,\cdots,q)$. Then

$$B_{\rho_0 a_0}(w_0)=q!(-1)^{\frac{q(q-1)}{2}}\omega^1\wedge\cdots\wedge\omega^q$$

and so, by (4.22)

$$(\theta_q \mid Y)(w_0)=(-1)^{\frac{q(q-1)}{2}}\left(\frac{\sqrt{-1}}{2\pi}\right)^q\omega^1\wedge\cdots\wedge\omega^q\wedge\bar{\omega}^1\wedge\cdots\wedge\bar{\omega}^q$$
$$+\text{(non-negative terms)}.$$

Thus $\theta_q \mid Y$ is a volume element. Q. E. D.

Remark. We can generalize Theorem 4.2′ a little as follows: Let $Y\subset W$ be an irreducible subvariety of dimension s, and let $q_1,\cdots,q_t$, $1\leqq q_\mu\leqq r$, satisfy $\sum_{\mu=1}^{t}q_\mu=s$. Then we have

THEOREM 4.3. $\langle c_{q_1}\cdots c_{q_t}, y\rangle>0$.

Proof. For simplicity, we assume that $t=2$. Thus let p, q satisfy $1\leqq p,q\leqq r$ and $p+q=s$. We agree on the ranges of indices $1\leqq\alpha,\beta\leqq p$; $p+1\leqq\mu,\nu\leqq p+q$. Then, by (4.22),

$$\theta_p\wedge\theta_q=\frac{(-1)^{\frac{1}{2}(p(p-1)+q(q-1))}}{p!q!}\left(\frac{\sqrt{-1}}{2\pi}\right)^{p+q}\sum_{\substack{\rho=(\rho_1,\cdots,\rho_p)\\ a=(a_1,\cdots,a_p)\\ \sigma=(\sigma_1,\cdots,\sigma_q)\\ d=(d_1,\cdots,d_q)}}$$
$$\{B_{\rho a}\wedge\bar{B}_{\rho a}\wedge B_{\sigma d}\wedge\bar{B}_{\sigma d}\}$$

$$=\frac{(-1)^{\frac{1}{2}(p+q)(p+q-1)}}{p!q!}\left(\frac{\sqrt{-1}}{2\pi}\right)^{p+q}\sum_{\substack{\rho,\sigma\\ a,d}}B_{\rho a}\wedge B_{\sigma d}\wedge\overline{B_{\rho a}\wedge B_{\sigma d}}.$$

As above, we let Y pass through w, and be given locally by $z^j=0$ $(s<j\leqq n+r)$. Furthermore, we choose our local $\boldsymbol{E}$-frame so that $b_{1\rho j}(w_0)=0$ for $\rho>p, j>p$, $b_{1\alpha\beta}(w_0)=\delta_\beta{}^\alpha$; and $b_{2\rho j}(w_0)=0$ for $0\leqq j<p, \rho>q$, $j>p+q$, $b_{2\mu-p\nu}(w_0)=\delta_\nu{}^\mu$. It is possible to do this since $\boldsymbol{E}\rightarrow W$ is ample. Now let $\omega^1,\cdots,\omega^{n+r}$ be chosen so that $\omega^j(w_0)=dz^j$ and $\omega^j=0$ on Y for $j>p+q$. Then, if $\rho_0=(1,\cdots,p)$, $\sigma_0=(p+1,\cdots,p+q)$, $a_0=(1,\cdots,1)$, $d_0=(2,\cdots,2)$,

$$\frac{1}{p!q!}(B_{\rho_0 a_0}\wedge B_{\sigma_0 d_0}\mid Y)(w_0)=(-1)^{\frac{1}{2}(p+q)(p+q-1)}$$
$$\omega^1\wedge\cdots\wedge\omega^p\wedge\omega^{p+1}\wedge\cdots\wedge\omega^{p+q}.$$

Thus $\theta_p \wedge \theta_q \mid Y$ is a volume element. Q. E. D.

Appendix to § IV: Numerically Positive Bundles. Let W be an algebraic manifold of dimension n.

Definition. Let ϕ be a real closed form of type (s,s) on W. Then we write $\phi > 0$ to mean that $\int_Y \phi > 0$ for all subvarieties $Y \subset W$ of dimension s.

Now let $E \to W$ be a holomorphic bundle with total Chern class $c(E) = \sum_{i=0}^{r} c_i$ $(c_i \in H^{2i}(W, Z))$. We consider vectors $I = (i_1, \cdots, i_r)$ of non-negative integers and set $c_I = c_1^{i_1} \cdots c_r^{i_r}$, $|I| = i_1 + 2i_2 + \cdots + ri_r$. Then Theorem 4.3 may be rephrased as follows: If $E \to W$ is ample, then

$$c_I > 0 \text{ for all } I \text{ with } |I| \leqq n. \tag{A.1}$$

As remarked in IV. 3, (A. 1) characterizes positive line bundles but for general vector bundles there are additional more subtle numerical inequalities which hold for ample bundles. After discussing these in §§ A. 1 and A. 2, we shall in § A. 3 outline a proof of the Theorem that, if all the numerical inequalities for $E \to W$ are satisfied, then the bundles E^μ are ample for $\mu \geqq \mu_0$.

A. 1. The Cone of Positive Polynomials for Ample Bundles. As an example of the additional numerical relations which hold for general bundles, we show

PROPOSITION A. 1. *If $E \to W$ is ample, then*

$$c_1^2 - 2c_2 > 0. \tag{A.2}$$

Proof. We use the notations from IV. 4. Thus

$$\theta_1 = \frac{\sqrt{-1}}{2\pi} \sum_{\substack{1 \leqq a \leqq m \\ 1 \leqq \rho \leqq r}} b_{a\rho} \wedge \bar{b}_{a\rho}$$

and

$$\theta_2 = -\tfrac{1}{2}\left(\frac{\sqrt{-1}}{2\pi}\right)^2 \sum_{\substack{\rho < \sigma \\ 1 \leqq a, d \leqq m}} b_{a\rho} \wedge b_{d\sigma} \wedge (\overline{b_{a\rho} \wedge b_{d\sigma} + b_{d\rho} \wedge b_{a\sigma}}).$$

Consequently,

$$\theta_1^2 - 2\theta_2$$

$$= -\left(\frac{\sqrt{-1}}{2\pi}\right)^2 \{\sum b_{a\rho} \wedge b_{d\sigma} \wedge \bar{b}_{a\rho} \wedge \bar{b}_{d\sigma} - (\sum_{\rho<\sigma} b_{a\rho} \wedge b_{d\sigma} \wedge \bar{b}_{a\rho} \wedge \bar{b}_{d\sigma} + b_{a\rho} \wedge b_{d\sigma} \wedge \bar{b}_{d\rho} \wedge \bar{b}_{a\sigma})\}.$$

Now

$$0 \leqq \tfrac{1}{2} \sum_{a,d} (b_{a\rho} \wedge b_{d\sigma} - b_{d\rho} \wedge b_{a\sigma}) \wedge \overline{(b_{a\rho} \wedge b_{d\sigma} - b_{d\rho} \wedge b_{a\sigma})}$$
$$= \sum_{a,d} b_{a\rho} \wedge b_{d\sigma} \wedge \bar{b}_{a\rho} \wedge \bar{b}_{d\sigma} - \sum_{a,d} b_{a\rho} \wedge b_{d\sigma} \wedge \bar{b}_{d\rho} \wedge \bar{b}_{a\sigma}.$$

Adding this last inequality to $\theta_1{}^2 - 2\theta_2$, we get

$$\theta_1{}^2 - 2\theta_2$$
$$\geqq -\left(\frac{\sqrt{-1}}{2\pi}\right)^2 \{ \sum_{\rho,\sigma} b_{a\rho} \wedge b_{d\sigma} \wedge \bar{b}_{a\rho} \wedge \bar{b}_{d\sigma} - 2 \sum_{\rho<\sigma} b_{a\rho} \wedge b_{d\sigma} \wedge \bar{b}_{a\rho} \wedge \bar{b}_{d\sigma} \}$$
$$= -\left(\frac{\sqrt{-1}}{2\pi}\right)^2 \sum b_{a\rho} \wedge b_{d\rho} \wedge \bar{b}_{a\rho} \wedge \bar{b}_{d\rho}.$$

As in the proof of Theorem 4. 3, we find easily that, if $\boldsymbol{E} \to W$ is ample, then

$$-\left(\frac{\sqrt{-1}}{2\pi}\right)^2 \sum b_{a\rho} \wedge b_{d\rho} \wedge \bar{b}_{a\rho} \wedge \bar{b}_{d\rho} > 0.$$

Q. E. D.

To get an idea what Proposition A. 1 means, consider the case where $\boldsymbol{E} = \boldsymbol{F} \oplus \boldsymbol{G}$ is a sum of positive line bundles $\boldsymbol{F}$, $\boldsymbol{G}$ with respective Chern classes f, g. Then $c_1{}^2 - 2c_2 = (f+g)^2 - 2fg = f^2 + g^2 > 0$.

In this particular case with $\boldsymbol{E} = \boldsymbol{F} \oplus \boldsymbol{G}$, we have $c_1{}^2 - \beta c_2 > c_1{}^2 - 4c_2 = (f-g)^2 \geqq 0$ provided that $0 \leqq \beta < 4$. However this stronger type of relation will not hold in general. In fact, one may find an ample bundle $\boldsymbol{E}$ for which we have an exact sequence $0 \to \boldsymbol{F} \to \boldsymbol{E} \to \boldsymbol{G} \to 0$ such that:

(i) $\boldsymbol{G} \to W$ is positive;

(ii) $f \geqq 0$ and $f^2 = 0$; and

(iii) $fg = g^2$.

Then $c_1{}^2 - 3c_2 = (f+g)^2 - 3fg = -fg + g^2 = 0$.

This all suggests the following construction: Let $x_1, \cdots, x_r$ be indeterminants (over $\boldsymbol{Q}$) and write formally

$$1 + x_1 t + \cdots + x_r t^r = (1 + \gamma_1 t) \cdots (1 + \gamma_r t).$$

We consider the ring $\boldsymbol{R}$ of polynomials

$$P(x_1, \cdots, x_r) = \sum_I p_I x^I$$
$$(I = (i_1, \cdots, i_r), x^I = x_1{}^{i_1} \cdots x_r{}^{i_r}, p_I \in \boldsymbol{Q}),$$

and we grade $\boldsymbol{R}$ by assigning to x^I the weight $|I| = i_1 + 2i_2 + \cdots + ri_r$. Then $\boldsymbol{R} = \bigoplus_{s \geqq 0} \boldsymbol{R}^s$ and $\boldsymbol{R}^s \cdot \boldsymbol{R}^t \subseteq \boldsymbol{R}^{s+t}$. We want to define the *positive cone*

$\mathbf{\Pi}\subset \boldsymbol{R}$ and this is done as follows: An element $P\in \boldsymbol{R}^s$ may be uniquely written as a homogeneous symetric polynomial $\sigma(\gamma_1,\cdots,\gamma_r)$ of degree s:

$$P(x_1,\ \cdot\cdot,x_r)=\sigma(\gamma_1,\cdots,\gamma_r)=\sum_{|J|=s}\sigma_J\gamma^J$$
$$(J=(j_1,\cdots,j_r),\gamma^J=\gamma_1^{j_1}\cdots\gamma_r^{j_r},|J|=j_1+\cdots+j_r).$$

We say then that P *is positive*, written $P>0$, if $P\neq 0$ and all $\sigma_J\geqq 0$. Let $\mathbf{\Pi}^s\subset \boldsymbol{R}^s$ be the set of positive polynomials, $\mathbf{\Pi}=\oplus\ \mathbf{\Pi}^s$ the *positive cone.*

Definition. We say that $\boldsymbol{E}\to W$ *lies in the position cone* if $P_s(c_1,\cdots,c_r)>0$ for all $P_s\in\mathbf{\Pi}^s$ $(0\leqq s\leqq n)$.

Remarks and Examples. (i) If $r=1$, $\mathbf{\Pi}^s=$ all polynomials λx^s with $\lambda>0$. Thus, if $\boldsymbol{E}\to W$ is a line bundle which lies in the positive cone, then $\boldsymbol{E}$ is positive and viceversa.

(ii) $\mathbf{\Pi}^1=$ polynomials λx_1 with $\lambda>0$, $\mathbf{\Pi}^2=$ polynomials $\alpha x_1^2+\beta x_2$ with $\alpha\geqq 0$, $\beta+2\alpha\geqq 0$, and where at least one of the inequalities is strict.

(iii) $\mathbf{\Pi}^s\cdot\mathbf{\Pi}^t\subset\mathbf{\Pi}^{s+t}$. Therefore $x^I\in\mathbf{\Pi}^{|I|}$ since

$$x^I=x_1^{i_1}\cdots x_r^{i_r}\text{ and }x_q=\sum_{1\leqq j_1<\cdots<j_q\leqq r}\gamma_{j_1}\cdots\gamma_{j_q}\in\mathbf{\Pi}^q.$$

(iv) Given r, there is for each s an important distinguished element $\Phi_s\in\boldsymbol{R}^s$ which arises from the positive symmetric polynomial

$$\sigma_s(\gamma_1,\cdots,\gamma_r)=\sum_{j_1+\cdots+j_r=s}\gamma_1^{j_1}\cdots\gamma_r^{j_r}. \tag{A.3}$$

For example,

$$\begin{cases}\Phi_0(x_1,\cdots,x_r)=1\\ \Phi_1(x_1,\cdots,x_r)=x_1\\ \Phi_2(x_1,\cdots,x_r)=x_1^2-x_2\\ \Phi_3(x_1,\cdots,x_r)=x_1^3-2x_1x_2+x_3,\ \text{etc.}\end{cases} \tag{A.4}$$

An interpretation of the polynomials Φ_s will be given in Lemma A.1 in § A.3 below.

(v) The same sort of calculations as in the proofs of Theorems 4.2, 4.3 and Proposition A.1 gives us

THEOREM A.1. *If* $\boldsymbol{E}\to W$ *is ample, then* $\boldsymbol{E}$ *lies in the positive cone.*

(vi) The condition that $\boldsymbol{E}\to W$ lie in the positive cone is, in some respects, rather weak. For example, let X, Y be non-singular curves and $\boldsymbol{G}\to X$, $\boldsymbol{H}\to Y$ positive line bundles. Set $W=X\times Y$, $\boldsymbol{F}=\boldsymbol{G}\otimes\boldsymbol{H}$, $\boldsymbol{E}=\boldsymbol{F}\oplus\boldsymbol{H}$. Then every symmetric power $\boldsymbol{E}^\mu$ contains $\boldsymbol{H}^\mu$ as a direct summand, so that no power of E is ample. On the other hand, $\boldsymbol{E}$ is in the positive cone since $\alpha c_1^2+\beta c_2=\alpha(f+h)^2+\beta fh=(2\alpha+\beta)gh>0$.

A. 2. Numerically Positive Bundles. The example (vi) in § A. 1 shows that there are still further relations which should be present for an ample bundle $\boldsymbol{E}\to W$. To see what these should be, we consider the following two fundamental properties of an ample (or positive, or weakly positive) bundle: (1) If $\boldsymbol{E}\to W$ is ample and if $Y\subset W$ is a subvariety, then $\boldsymbol{E}_Y=\boldsymbol{E}\mid Y$ is ample; and (2) if $\boldsymbol{E}\to W$ is ample, then any quotient bundle $\boldsymbol{Q}$ of $\boldsymbol{E}$ is ample.

Definition. Let $\boldsymbol{E}\to W$ be a holomorphic vector bundle. Then we shall say that $\boldsymbol{E}$ is *numerically positive* if, for any subvariety $Y\subset W$ and any quotient bundle $\boldsymbol{Q}$ of $\boldsymbol{E}_Y$, $\boldsymbol{Q}$ lies in the positive cone.

In particular, the bundle in example (vi) of A. 1 is not numerically positive since $\boldsymbol{H}$ is a quotient bundle of $\boldsymbol{E}$ and $\boldsymbol{H}\mid X\times$ (pt.) is trivial. In fact, our main result is

THEOREM IV. *Let $\boldsymbol{E}\to W$ be a holomorphic vector bundle. Then, if $\boldsymbol{E}$ is positive, the bundles $\boldsymbol{E}^\mu$ are numerically positive if $\mu\geqq\mu_0$. Conversely, if $\boldsymbol{E}$ is numerically positive, then the bundles $\boldsymbol{E}^\nu$ are positive for $\nu\geqq\nu_0$.*

Before giving the proof of Theorem IV in § A. 3 below, we make some preliminary remarks. First, if $\boldsymbol{E}\to W$ is positive, then the bundles $\boldsymbol{E}^\mu$ $(\mu\geqq\mu_0)$ are ample for $\mu\geqq\mu_0$ so that one-half of the Theorem follows from Theorem A. 1. Also, the bundles $\boldsymbol{E}^\nu$ $(\nu\geqq\nu_0)$ are all positive if, and only if, the following vanishing theorem holds: For $\mathfrak{S}\to W$ a coherent analytic sheaf, we have

$$H^q(W,\mathfrak{S}\otimes\Omega(\boldsymbol{E}^\nu))=0 \text{ for } q\geqq 1 \text{ and } \nu\geqq\nu_0(\mathfrak{S}). \tag{A.5}$$

To prove the other half of Theorem IV, we shall thus prove (A. 5).

A. 3. Proof of Theorem IV.

(i) Let $\boldsymbol{E}^*\to W$ be the dual bundle to $\boldsymbol{E}$ and $P(\boldsymbol{E}^*)\xrightarrow{\pi}W$ the associated bundle of projective spaces. Over $P(\boldsymbol{E}^*)$ there is a tautological line bundle $\boldsymbol{L}^*$ whose fibre $\boldsymbol{L}^*_\xi$ $(\xi\in P(\boldsymbol{E}^*))$ is the line $[\xi]\subset\boldsymbol{E}^*_{\pi(\xi)}$. There is an obvious mapping $\psi\colon\boldsymbol{L}^*\to\pi^{-1}(\boldsymbol{E}^*)$ and we define $\phi\colon\pi^{-1}(\boldsymbol{E})\to\boldsymbol{L}$ by

$$\langle\phi(\xi,e),\lambda^*\rangle=\langle\psi(\xi,\lambda^*),e\rangle \qquad (\xi\in P(\boldsymbol{E}^*),$$

$e\in\boldsymbol{E}_{\pi(\xi)},\lambda^*\in\boldsymbol{L}^*_\xi)$. Then ϕ induces $\phi_\mu\colon\pi^{-1}(\boldsymbol{E}^\mu)\to\boldsymbol{L}^\mu$ and, for any coherent sheaf $\mathfrak{S}\to W$, the cohomology mappings

$$H^q(W, \mathfrak{S} \otimes \Omega(\boldsymbol{E}^\mu)) \to H^q(P(\boldsymbol{E}^*), \pi^{-1}(\mathfrak{S} \otimes \Omega(\boldsymbol{E}^\mu))) \xrightarrow{\phi_\mu} H^q(P(\boldsymbol{E}^*), \pi^{-1}(\mathfrak{S}) \otimes \Omega(\boldsymbol{L}^\mu)) \tag{A.6}$$

are all isomorphisms. Thus, by (A. 5), in order to prove Theorem IV it will suffice to show

PROPOSITION A. 2. *If* $\boldsymbol{E} \to W$ is numerically positive, then $\boldsymbol{L} \to P(\boldsymbol{E}^*)$ is numerically positive.

(ii) We suppose that we have in $\boldsymbol{E} \to W$ an Hermitian metric and let θ_q $(0 \leqq q \leqq r)$ be the real (q, q) form which represents $c_q(\boldsymbol{E}) \in H^{2q}(W, \boldsymbol{Z})$ (cf. (4. 17)). Then there is naturally induced in $\boldsymbol{L} \to P(\boldsymbol{E}^*)$ a metric and we denote by θ the corresponding $(1,1)$ form on $P(\boldsymbol{E}^*)$ which represents $c_1(\boldsymbol{L}) = \omega$.

LEMMA A. 1. $\pi_*(\theta^{s+r-1}) = \Phi_s(\theta_1, \cdots, \theta_r) +$ *(exact form) where* π_* *is integration over the fibre and where* $\Phi_s \in \mathbf{\Pi}^s$ *is given by* (A. 3).

Proof. Over $P(\boldsymbol{E}^*)$ *we have the exact bundle sequence*

$$0 \to \boldsymbol{F} \to \pi^{-1}(\boldsymbol{E}) \to \boldsymbol{L} \to 0. \tag{A.7}$$

Therefore, if $c(\boldsymbol{F}) = d_0 + \cdots + d_{r-1}$, we have

$$c_q = d_q + d_{q-1}\omega \qquad (0 \leqq q \leqq r). \tag{A.8}$$

By (A. 8), the alternating sum $\sum_{q=0}^{r} (-1)^q c_q \omega^{r-q}$ telescopes to zero, and we get the identity $\omega^r = \sum_{q>0} (-1)^{q+1} c_q \omega^{r-q}$. More generally, we have

$$\omega^{s+r} = \sum_{r \geqq q > 0} (-1)^{q+1} c_q \omega^{s+r-q}. \tag{A.9}$$

Since $\theta^{r-1} \mid \pi^{-1}(w)$ is a normalized volume, $\pi_*(\theta^{r-1})$ is the constant function 1 on W. Inductively we assume that

$$\pi_*(\theta^{s+r-t}) = \Phi_{s+1-t}(\theta_1, \cdots, \theta_r) + \text{(exact form)} \quad (1 \leqq t \leqq s + r) \tag{A.10}$$

and we shall show that

$$\pi_*(\theta^{s+r}) = \Phi_{s+1}(\theta_1, \cdots, \theta_q) + \text{(exact form)}. \tag{A.11}$$

Since (by (A. 9))

$$\theta^{s+r} = \sum_{0 < q \leqq r} (-1)^{q+1} \theta^{s+r-q} \wedge \pi^*(\theta_q) + \text{(exact form)},$$

$$\pi_*(\theta^{s+r}) = \sum_{0 < q \leqq r} (-1)^{q+1} \pi_*(\theta^{s+r-q}) \wedge \theta_q + \text{(exact form)}.$$

By (A.10), $\pi_*(\theta^{s+r}) = \sum_{0<q\leqq r} (-1)^{q+1}\theta_q \wedge \Phi_{s+1-q}(\theta_1, \cdots, \theta_q) +$ (exact form) and where $\Phi_s = 0$ for $s < 0$. Thus what we have to prove is

$$\Phi_{s+1}(x_1, \cdots, x_r) = \sum_{0\leqq q<r} (-1)^q x_{q+1}\Phi_{s-q}(x_1, \cdots, x_r). \tag{A.12}$$

By (A.3) and induction, (A.12) amounts to the identity

$$\begin{aligned}\sum_{i_1+\cdots+i_r=s+1} \gamma_1^{i_1} \cdots \gamma_r^{i_r} = \sum (-1)^q &\sum_{1\leqq j_1<\cdots<j_{q+1}\leqq r} \\ &\gamma_{j_1} \cdots \gamma_{j_{q+1}} \sum_{k_1+\cdots+k_r=s-q} \gamma_1^{k_1} \cdots \gamma_r^{k_r}.\end{aligned} \tag{A.13}$$

We now use the notations

$$I = (i_1, \cdots, i_r),\ K = (k_1, \cdots, k_r),\ |I| = i_1 + \cdots + i_r,\\ \gamma^I = \gamma_1^{i_1} \cdots \gamma_r^{i_r}.$$

Also, A, B, C will stand for r-tuples $(a_1, \cdots, a_r)$ with $a_j = 0$ or 1. With this understood, (A.13) becomes

$$\gamma^I = \sum_{A+K=I} (-1)^{|A|-1}\gamma^A\gamma^K. \tag{A.14}$$

Given $I = (i_1, \cdots, i_r)$, we define $A_I = (a_1, \cdots, a_r)$ by $a_j = 0$ if $i_j = 0$, $a_j = 1$ if $i_j > 0$. Then

$$\sum_{A+K=I} (-1)^{|A|-1}\gamma^A\gamma^K = \Big(\sum_{B+C=A_I} (-1)^{|B|-1}\gamma^B\gamma^C\Big)\gamma^{I-A_I}. \tag{A.15}$$

Since $\gamma^I = \gamma^{A_I}\gamma^{I-A}$, by (A.14) and (A.15) it will suffice to prove

$$\gamma^{A_I} = \sum_{B+C=A_I} (-1)^{|B|-1}\gamma^B\gamma^C.$$

By changing the notation, this becomes the equation

$$\begin{aligned}\gamma_1 \cdots \gamma_t = \sum_i \gamma_i(\gamma_1 \cdots \hat{\gamma}_i \cdots \gamma_t) - \sum_{i<j} \gamma_i\gamma_j(\gamma_1 \cdots \hat{\gamma}_i \cdots \hat{\gamma}_j \cdots \gamma_t) \\ + \cdots + (-1)^{t-1}\gamma_1 \cdots \gamma_t.\end{aligned} \tag{A.16}$$

Now (A.16) is easy to check and this proves (A.11).

(iii) We pause to give an application of Lemma A.1.

For $Y \subset W$ a subvariety and $\boldsymbol{F} \to Y$ a line bundle, set

$$\chi(Y, \boldsymbol{F}(\mu)) = \sum_{0\leqq q\leqq s=\dim Y} (-1)^q H^q(Y, \Omega(\boldsymbol{F} \otimes \boldsymbol{E}^\mu)).$$

Furthermore we set $\chi(Y, \mu) = \chi(Y, \boldsymbol{l}(\mu))$ where $\boldsymbol{l}$ is the trivial bundle.

PROPOSITION A.3. *$\chi(Y, \mu)$ is a polynomial in μ with leading coefficient*

$$\frac{\mu^{s+r-1}}{(s+r-1)\,!}\int_Y \Phi_s(\theta_1,\cdots,\theta_q) \tag{A.17}$$

Proof. Set $Y_\pi = \pi^{-1}(Y) \subset P(\boldsymbol{E}^*)$ and

$$\chi(Y_\pi,\mu) = \sum_{q=0}^{s+r-1}(-1)^q \dim H^q(Y_\pi,\Omega(\boldsymbol{L}^\mu)).$$

By (A.6) with $\mathfrak{S} = \mathfrak{O}_Y$, $\chi(Y,\mu) = \chi(Y_\pi,\mu)$. By the Riemann-Roch theorem ([9], §20.4), $X(Y_\pi,\mu)$ is a polynomial in μ with leading coefficient

$$\frac{\mu^{s+r-1}}{(s+r-1)\,!}\langle\omega^{s+r-1},Y_\pi\rangle = \frac{\mu^{s+r-1}}{(s+r-1)\,!}\int_{Y_\pi}\theta^{s+r-1}. \tag{A.18}$$

But $\int_Y \theta^{s+r-1} = \int_Y \pi_*(\theta^{s+r-1})$ and then (A.17) follows from this, (A.18), and (A.11).

(iv) *Definition.* A connected non-singular subvariety $Z \subset P(\boldsymbol{E}^*)$ is said to be *fibered over* $Y \subset W$ if $\pi(Z) = Y$, if $\pi \mid Z$ has constant rank, and if the varieties $Z \cap \pi^{-1}(y) \subset P(\boldsymbol{E}^*_y)$ all have degree 1.

LEMMA. A.2. *The subvarieties $Z \subset P(\boldsymbol{E}^*)$ which are fibered over $Y \subset W$ are in a one-to-one correspondence with the quotient bundles $\boldsymbol{Q}_Z$ of $\boldsymbol{E}_Y = \boldsymbol{E} \mid Y$.*

Proof. A point $\xi \in P(\boldsymbol{E}^*)$ is given by a line $[\xi] \subset \boldsymbol{E}^*_{\pi(\xi)}$. Let Z be fibered over Y and let $y \in Y$. We set $\boldsymbol{F}_y = \{e \in \boldsymbol{E}_y \colon \langle[\xi],e\rangle = 0$ for all $\xi \in Z \cap \pi^{-1}(y)\}$. Then $\bigcup_{y\in Y}\boldsymbol{F}_y = \boldsymbol{F}_Z$ form a vector bundle over Y which is clearly a sub-bundle of $\boldsymbol{E}_Y$. We then set $\boldsymbol{Q}_Z = \boldsymbol{E}_Y/\boldsymbol{F}_Z$. Since this process is reversible, we are done.

Remark. In a natural way, $Z \cong P(\boldsymbol{Q}^*_Z)$ and $\boldsymbol{L} \mid Z$ corresponds to the dual of the tautological line bundle $\boldsymbol{L}^*_Z$ over $P(\boldsymbol{Q}^*_Z)$.

(v) LEMMA A.3. *If $Z \subset P(\boldsymbol{E}^*)$ is fibered over $Y \subset W$ and if $t = \dim Z$, then, if $\boldsymbol{E} \to W$ is numerically positive,*

$$\int_Z \theta^t > 0. \tag{A.19}$$

Proof. Let ψ be a real (1,1) form on Z which represents $c_1(\boldsymbol{L}_Z)$. Assume that $s = \dim Y$ so that $\boldsymbol{Q}_Z$ has fibre dimension $t-s+1$ and write $c(\boldsymbol{Q}_Z) = h_0 + \cdots + h_{t-s+1}$. Then, by the Remark following Lemma A.2 and Lemma A.1,

$$\int_Z \theta^t = \int_Z \psi^t = \int_Y \Phi_s(h_0,\cdots,h_{t-s+1}). \tag{A.20}$$

Since $\boldsymbol{E} \to W$ is numerically positive, $\boldsymbol{Q}_Z \to Y$ is in the positive cone and $\int_Y \Phi_s(h_0, \cdots, h_{t-s+1}) > 0$. Q. E. D.

(vi) Now in order to prove Proposition A. 2 we must show that, for any irreducible $Z \subset P(\boldsymbol{E}^*)$ of dimension t,

$$\int_Z \theta^t > 0. \tag{A.21}$$

For those Z which are fibered over some $Y \subset W$, (A. 19) gives (A. 21). In fact, Lemmas A. 2 and A. 3 contain the essential idea of the general proof but they need to be extended a little. In order to illustrate what has to be done, we prove a special case of (A. 21).

Lemma A. 4. *(A. 21) holds if $t = 1$.*

Proof. If $\pi(Z) =$ point we are done so that we may assume that Z is an m-sheeted branched covering over $Y = \pi(Z)$. Let $\Delta \subset Z$ be the branch locus and $D = \pi(\Delta)$. For $y \in Y$, let $z_1(y), \cdots, z_m(y)$ be the (not necessarily distinct) points in $\pi^{-1}(y) \cap Z$ and set $\boldsymbol{L}^{\#}{}_y = \boldsymbol{L}_{z_1(y)} \otimes \cdots \otimes \boldsymbol{L}_{z_m(y)}$. Then $\boldsymbol{L}^{\#} = \bigcup_{y \in Y} \boldsymbol{L}^{\#}{}_y$ is a line bundle over Y and there is an obvious mapping $\phi : \boldsymbol{E}_Y{}^m \to \boldsymbol{L}^{\#} \to 0$. Since $\boldsymbol{E} \to W$ is numerically positive, $\int_Y \psi > 0$ where $\psi = c_1(\boldsymbol{L}^{\#})$. Now $\int_Z \theta = \int_Y \pi_* \theta = \int_Z \pi^*(\pi_* \theta)$. But

$$(\pi_* \theta)(y) = \frac{1}{m} \sum_{i=1}^{m} \theta(z_i(y)) \text{ and so } \pi^*(\pi_* \theta) = \frac{1}{m} c_1(\pi^{-1} \boldsymbol{L}^{\#})$$

since $\theta = c_1(\boldsymbol{L})$. Thus $\int_Z \theta = \frac{1}{m} \int_Z \pi^* \psi > 0$. Q. E. D.

(vii) For a variety V, let $\boldsymbol{H}_*(V) = \sum \boldsymbol{H}_q(V)$ be the rational algebraic homology ring: $\boldsymbol{H}_q(V) \subset \boldsymbol{H}_{2q}(V, \boldsymbol{Q})$ is the subspace over $\boldsymbol{Q}$ generated by the algebraic cycles. Let $\gamma_1, \cdots, \gamma_k$ be a minimal set of effective * generators for $\boldsymbol{H}_*(W)$ and set $\xi_\alpha = \pi^{-1}(\gamma_\alpha)$ $(1 \leqq \alpha \leqq k)$. If ω is an algebraic cycle on $P(\boldsymbol{E}^*)$ dual to $c_1(\boldsymbol{L})$, then $\boldsymbol{H}_*(P(\boldsymbol{E}^*))$ is generated by the products $\phi_{j\alpha} = \omega^j \cdot \xi_\alpha$ with the relation $\sum_{q=0}^{r} (-1)^q \omega^{r-q} \cdot \pi^{-1}(c_q) = 0$ $(c_q \in \boldsymbol{H}_q(W))$.

Lemma A. 6. *ω may be chosen to be an effective cycle.*

Proof. We must show that $H^0(P(\boldsymbol{E}^*), \Omega(\boldsymbol{L}^\mu)) \neq 0$ for μ sufficiently large, i.e. $H^0(W,\Omega(\boldsymbol{E}^\mu))\neq 0$ for μ large. Let $H\subset W$ be a hyperplane section. By induction on the dimension of W, we may assume that the bundles $\boldsymbol{E}^\nu{}_H$ $(\nu\geqq\nu_0, \boldsymbol{E}_H=\boldsymbol{E}\mid H)$ are ample. Set $\boldsymbol{D}=[H]$ so that $\Omega(\boldsymbol{D}^*)$ is the ideal sheaf of $H\subset W$. Then we may find λ_0 so that

---

* i. e. $\gamma_j = \sum \eta_{j\gamma} Y_\gamma$ where Y_γ are irreducible subvarieties and $\eta_{j\gamma} > 0$.

$$H^q(H,\Omega(\boldsymbol{D}^\sigma{}_H\otimes\boldsymbol{E}^\lambda{}_H))=0$$

for $\sigma\geqq 0$, $\lambda\geqq\lambda_0$, $0<q$. From the exact cohomology sequences of

$$0\to\Omega(\boldsymbol{D}^{\sigma-1}\otimes\boldsymbol{E}^\lambda)\to\Omega(\boldsymbol{D}^\sigma\otimes\boldsymbol{E}^\lambda)\to\Omega(\boldsymbol{D}^\sigma{}_H\otimes\boldsymbol{E}^\lambda{}_H)\to 0$$

we get $H^q(W,\Omega(\boldsymbol{D}^\sigma\otimes\boldsymbol{E}^\lambda))=0$ for $\sigma\geqq 0$, $\lambda\geqq\lambda_0$, $q>1$ and

$$H^1(W,\Omega(\boldsymbol{D}^{\sigma-1}\otimes\boldsymbol{E}^\lambda))\to H^1(W,\Omega(\boldsymbol{D}^\sigma\otimes\boldsymbol{E}^\lambda))\to 0$$

for $\sigma\geqq 0$, $\lambda\geqq\lambda_0$. Since $\chi(W,\lambda)=\sum(-1)^q\dim H^q(W,\Omega(\boldsymbol{E}^\lambda))$ is a polynomial in λ with leading coefficient $c\lambda^{n+r-1}$ $(c>0)$, it follows that $\dim H^0(W,\Omega(\boldsymbol{E}^\lambda))\to\infty$ as $\lambda\to\infty$. This proves the lemma.

(viii) We shall call a variety $Z\subset P(\boldsymbol{E}^*)$ *linear* if, for each $w\in W$, the irreducible components of $Z\cap\pi^{-1}(w)$ are each linear subspaces of $P(\boldsymbol{E}^*{}_w)$. Since $\phi_{j\alpha}=\omega^j\cdot\pi^{-1}(\gamma_\alpha)$ ω is effective, and $\omega\mid P(\boldsymbol{E}^*{}_w)$ is a hyperplane section, it follows that the varieties $\phi_{j\alpha}$ $(0\leqq j\leqq r-1)$ are linear. The same is true of any variety $Y_\pi\cdot\omega^j$ $(0\leqq j\leqq r-1)$, $Y\subset W$, $Y_\pi=\pi^{-1}(Y)$.

Let now $Z\subset P(\boldsymbol{E}^*)$ be irreducible of dimension $t-1$ and suppose that $Y=\pi(Z)$ is of dimension s. Then $Z\cdot\pi^{-1}(y)$ $(y\in Y$ generic) is rationally equivalent to $\nu(Z)\lambda_y{}^{t-s}$ where $\nu(Z)>0$ and $\lambda_y{}^{t-s}$ is a linear subspace of dimension $t-s$ in $P(\boldsymbol{E}^*{}_y)$. Suppose that $Y\sim\sum\rho_\alpha\gamma_\alpha$ so that $Y_\pi\sim\sum\rho_\alpha\xi_\alpha$. Then $Z\sim\nu(Z)\sum\rho_\alpha\phi_{(r-1)-(t-s)}\sim\nu(Z)\omega^{(r-1)-(t-s)}\cdot Y_\pi$. The conclusion is

LEMMA A.7. *It will suffice to verify* (A.21) *for linear varieties* $Z\subset P(\boldsymbol{E}^*)$.

LEMMA A.8. (A.21) *is true for linear varieties.*

Proof. Let $Z\subset P(\boldsymbol{E}^*)$ be irreducible and linear. We shall assume that $Z_y=Z\cdot P(\boldsymbol{E}^*{}_y)$ is connected $(y\in Y=\pi(Z))$, the general case may be deduced from this special case by using branched coverings and symmetric products as in Lemma A.5. Now Z_y spans a linear subspace $\boldsymbol{F}_y\subset\boldsymbol{E}^*{}_y$ and $\boldsymbol{F}=\bigcup_{y\in Y}\boldsymbol{F}_y$ is a so-called quasi-linear space [6]. This $\boldsymbol{F}$ defines (in the usual manner) a coherent quotient sheaf $\mathfrak{Q}$ of $\Omega(\boldsymbol{E}_Y)$ $(\mathfrak{Q}=\Omega(\boldsymbol{E}_Y)/\mathfrak{S}$ where the sections of $\mathfrak{S}\subset\Omega(\boldsymbol{E})$ are those sections s of $\Omega(\boldsymbol{E}_Y)$ such that $\langle\boldsymbol{F}_y,s(y)\rangle=0$ for all y in the domain of s). The proof now may be completed by using the Riemann-Roch and directly generalizing the argument of Lemma A. 3 (which applied to the case where $\mathfrak{Q}$ is locally free).

V. Positive Embeddings and Monoidal Transformations.

1. Monoidal transformations. Let $X \subset W$ be an embedding of a complex manifold X in a complex manifold W. Suppose that $\dim X = n$, $\operatorname{codim} X = r > 1$, and that $\boldsymbol{N} \xrightarrow{\pi} X$ is the normal bundle. By replacing each x in X by $P(\boldsymbol{N}_x)$, the projective normal space to X at x, we obtain a new manifold $\tilde{W}$, the *monoidal transform* of W along X. There is a hypersurface $\tilde{X} \subset \tilde{W}$ and a proper mapping $\tilde{\omega}\colon \tilde{W} \to W$ such that: (i) $\tilde{\omega}\colon \tilde{W} - \tilde{X} \cong W - X$; (ii) $\tilde{X}$ may be identified with $P(\boldsymbol{N})$ and $\tilde{\omega} \mid \tilde{X}$ is the map induced by π; and (iii) if $\boldsymbol{L} = [\tilde{X}]$, then $\boldsymbol{L} \mid P(\boldsymbol{N}_x) = \boldsymbol{H}^*$ where $\boldsymbol{H}$ is the hyperplane bundle.

Thus, e. g., if $\boldsymbol{E} \xrightarrow{\pi} X$ is a holomorphic vector bundle, we may embed $X \subset \boldsymbol{E}$ as zero cross section. Furthermore, $\boldsymbol{C}^*$ acts on $\boldsymbol{E} - X$ and gives rise to a line bundle $\boldsymbol{L}(\boldsymbol{E}) \to P(\boldsymbol{E})$. Using then the above notation, $\tilde{W} = \boldsymbol{L}(\boldsymbol{E})$, $\tilde{X} = P(\boldsymbol{E})$ embedded in $\tilde{W}$ as zero-section, and $\tilde{W} - \tilde{X}$ is naturally identified with $W - X$.

In order to describe monoidal transformations locally, we may assume that $W \subset \boldsymbol{C}^{n+r}$ is the unit polycylinder with coordinates $(z_1, \cdots, z_n; w_1, \cdots, w_r)$ and that $X \subset W$ is given by $w_1 = \cdots = w_r = 0$. Let P_{r-1} have homogeneous coordinates $[t_0, \cdots, t_r]$. Then $\tilde{W} \subset W \times P_{r-1}$ is given as

$$\{(z, w, t) : w_\rho t_\sigma - t_\rho w_\sigma = 0 ; 1 \leqq \rho, \sigma \leqq r\}.$$

The mapping $\tilde{\omega}\colon \tilde{W} \to W$ is induced by the projection $W \times P_{r-1} \to W$ and $\tilde{X} = \tilde{\omega}^{-1}(X) \cong X \times P_{r-1}$.

Finally, we shall yield to custom and write $\boldsymbol{E}^{-1}$ for $\boldsymbol{E}^*$ whenever $\boldsymbol{E}$ is a line bundle.

2. Construction of a blowing down map. We shall discuss the following question: Let $X \subset W$ be complex submanifold of codimension 1 and set $\boldsymbol{L} = [X]$. Suppose that $X \xrightarrow{\pi} Y$ is a fibre bundle with fibre $\pi^{-1}(y) = P_y \cong P_{r-1}$ and that $\boldsymbol{L} \mid P_y = \boldsymbol{H}^{-1}$. We should then like to show that there is a manifold $Z \supset Y$ such that W is the monoidal transform of Z along Y and where X is the inverse image of Y, the mapping $X \xrightarrow{\pi} Y$ being induced by the projection $\tilde{\omega}\colon W \to Z$. In this case we say that we can *blow down W along X.*

We shall prove that it is possible to blow down W along X in case W is compact and with the assumption that the (hypothetical) normal bundle of Y in Z is positive. More precisely, we let $\boldsymbol{L} \mid X = \boldsymbol{N}$ and assume that $\boldsymbol{N}$ has signature $(n, (r-1))$ (cf. § VI. 3 below) where $n + r - 1 = \dim X$. This assumption is an analytic way of expressing the geometric hypothesis that *no further blowing down of Z along Y can occur.*

THEOREM. *Let W be a closed algebraic variety and $X \subset W$ a subvariety which satisfies the above assumption. Then we can blow down W along X to a non-singular algebraic variety Z.*

Proof. (a) Let $\boldsymbol{F} \to W$ be a positive line bundle. Then $\boldsymbol{F} \mid P_y = \boldsymbol{H}^l$ and we assume that $l > 2$. Set $\boldsymbol{E} = \boldsymbol{F} \otimes \boldsymbol{L}^l \mid X$ and denote by $R^q{}_\pi(\Omega(\boldsymbol{E}))$ the q-th Leray sheaf of the pair $(\Omega(\boldsymbol{E}); X \xrightarrow{\pi} Y)$.

LEMMA 5.1. *There exists a line bundle $\boldsymbol{G} \to Y$ such that $\pi^{-1}(\boldsymbol{G}) = \boldsymbol{E}$.*

Proof. Let $\{U_\alpha\}$ be a covering of Y by small polycylinders. Then we may find isomorphisms

$$\phi_\alpha\colon \pi^{-1}(U_\alpha) \longrightarrow U_\alpha \times P_{r-1}\text{*} \quad \text{and} \quad \phi_\alpha \phi_\beta^{-1}\colon U_\beta \times P_{r-1} \to U_\alpha \times P_{r-1}$$

will be written as $\phi_\alpha \phi_\beta^{-1}(u_\beta, p_\beta) = (f_{\alpha\beta}(u_\beta), g_{\alpha\beta}(u_\beta) p_\beta)$ where $g_{\alpha\beta}(u_\beta)$ is a holomorphic family of projective transformations. Since $\boldsymbol{E} \mid P_y$ is trivial, it follows that $\boldsymbol{E} \mid \pi^{-1}(U_\alpha)$ is trivial and so we may choose an isomorphism

$$\psi_\alpha\colon \boldsymbol{E} \mid \pi^{-1}(U_\alpha) \longrightarrow \pi^{-1}(U_\alpha) \times \boldsymbol{C}.$$

Then we have

$$\eta_\alpha\colon \boldsymbol{E} \mid \pi^{-1}(U_\alpha) \longrightarrow U_\alpha \times P_{r-1} \times \boldsymbol{C}$$

where $\eta_\alpha = (\phi_\alpha, 1) \circ \psi_\alpha$. Clearly $\rho_{\alpha\beta} = \eta_\alpha \circ \eta_\beta^{-1}$ is given as

$$\rho_{\alpha\beta}(u_\beta, p_\beta, \xi_\beta) = (f_{\alpha\beta}(u_\beta),\, g_{\alpha\beta}(u_\beta) p_\beta,\, \gamma_{\alpha\beta}(u_\beta) \xi_\beta)$$

where $\gamma_{\alpha\beta}(u_\beta)$ is a no-where zero holomorphic function. It is then clear that, if we define $\boldsymbol{G} \to Y$ to be the line bundle given by the system of transition functions, $\{\gamma_{\alpha\beta}(u_\beta)\}$; then $\pi^{-1}(\boldsymbol{G}) = \boldsymbol{E}$.

(b) We recall now the formula

$$R_\pi{}^q(\mathfrak{S} \otimes_{\mathfrak{O}_X} \pi^{-1}\mathfrak{T}) \cong R_\pi{}^q(\mathfrak{S}) \otimes_{\mathfrak{O}_Y} \mathfrak{T} \tag{5.1}$$

where $\mathfrak{S} \to X$ and $\mathfrak{T} \to Y$ are locally free coherent sheaves.

---

* We remark that any fibre space with projective space as fibre is a fibre bundle.

LEMMA 5.2. *For any* $\mathfrak{X} \to Y$ *as above and any integer* μ, *there is a canonical isomorphism*

$$H^q(Y, \Omega(\boldsymbol{G}^{\mu}) \otimes \mathfrak{X}) \cong H^q(X, \Omega(\boldsymbol{E}^{\mu}) \otimes \pi^{-1}(\mathfrak{X}))$$

Proof. By Leray's theorem, there exists a spectral sequence $\{E_r\}$ such that $E_\infty \Rightarrow H^*(X, \Omega(\boldsymbol{E}^{\mu}) \otimes \pi^{-1}(\mathfrak{X}))$ and

$$E_2^{p,q} = H^p(Y, R_\pi{}^q(\Omega(\boldsymbol{E}^{\mu}) \otimes \pi^{-1}(\mathfrak{X}))).$$

Since $\boldsymbol{E} = \pi^{-1}(\boldsymbol{G})$,

$$R_\pi{}^q(\Omega(\boldsymbol{E}^{\mu}) \otimes_{\mathfrak{O}_X} \pi^{-1}(\mathfrak{X})) \cong R_\pi{}^q(\mathfrak{O}_X) \otimes_{\mathfrak{O}_Y} \Omega(\boldsymbol{G}^{\mu}) \otimes_{\mathfrak{O}_Y} \mathfrak{X}.$$

But clearly $R_\pi{}^q(\mathfrak{O}_X) = 0$ for $q > 0$ and $R_0{}^q(\mathfrak{O}_X) \cong \mathfrak{O}_Y$. Q.E.D.

(c) We now prove the

LEMMA 5.3. $\boldsymbol{G} \to Y$ *is a positive line bundle.*

Proof. By well-known techniques, it will suffice to show that for *any* $\mathfrak{X} \to Y$ a locally free coherent sheaf, $H^q(Y, \Omega(\boldsymbol{G}^{-\mu}) \otimes \mathfrak{X}) = 0$ for $0 \leqq q < n$ and $\mu \geqq \mu_0(\mathfrak{X})$. By the Lemma in (b), we may show that $H^q(X, \Omega(\boldsymbol{E}^{-\mu}) \otimes \mathfrak{S}) = 0$ for $0 \leqq q < n$ and $\mu \geqq \mu_0(\mathfrak{S})$, where $\mathfrak{S} \to X$ is coherent and locally free.

Now $\boldsymbol{E}^{-1} = (\boldsymbol{F} \mid X)^{-1} \otimes \boldsymbol{N}^{-l}$. Since $\boldsymbol{N}^{-l}$ has signature $((r-1), n)$ and $\boldsymbol{F}^{-1} \mid X$ has signature $(0, n+r-1))$, it follows that, at every point $x \in X$, the curvature form for $\boldsymbol{E}^{-1}$ has signature $((r-1-\alpha(x)), (n+\alpha(x)))$ where $\alpha(x) \geqq 0$. By the vanishing theorems (cf. § VII.3 below) it follows that $H^q(X, \Omega(\boldsymbol{E}^{-\mu}) \otimes \mathfrak{S}) = 0$ for $0 \leqq q < n$ and $\mu \geqq \mu_0(\mathfrak{S})$. Q.E.D.

(d) By replacing $\boldsymbol{G}$ by a multiple $\boldsymbol{G}^k$ if necessary, we may assume that $|\boldsymbol{G}|$ gives a projective embedding $\Psi: Y \to P_s$. In doing this we replace $\boldsymbol{F}$ by $\boldsymbol{F}^k$ and l by kl.

From the lemma in (b) there is an isomorphism $H^0(Y, \Omega(\boldsymbol{G})) \cong H^0(X, \Omega(\boldsymbol{E}))$ and thus a diagram

$$\begin{tikzcd} X \arrow[d, "\pi"'] \arrow[dr, "\Lambda"] & \\ Y \arrow[r, "\Psi"'] & P_s \end{tikzcd}$$

where Λ is given by $|\boldsymbol{E}|$. We want to extend Λ to $\Phi: W \to P_N \supset P_s$ so that Φ is biregular on $W - X$. In order to do this, we prove

LEMMA 5.4. *$H^1(X, \Omega(\boldsymbol{E}\otimes \boldsymbol{N}^{-t}))=0$ for $0 \leqq t \leqq l$.*

Proof. Set $\boldsymbol{E}_t=\boldsymbol{F}\otimes \boldsymbol{L}^l \otimes \boldsymbol{L}^{-t} \mid X$. Then $\boldsymbol{E}_0=\boldsymbol{E}$ and $\boldsymbol{E}_l=\boldsymbol{F} \mid X$. Let ω on Y be the curvature form for $\boldsymbol{G}\rightarrow Y$ (cf. § VI.5 below) and ψ on X the form for $\boldsymbol{F} \mid X$. Then ψ gives a positive definite Kähler metric $\sum \omega_i \overline{\omega}_i$ on X and ω lifted to X is real and positive semi-definite. Thus we may choose the $\omega_1, \cdots, \omega_{n+r-1}$ so that $\omega=\frac{\sqrt{-1}}{2}\sum c_i\omega_i \wedge \overline{\omega}_i$ where $c_i \geqq 0$. Since $\boldsymbol{E}_l=\boldsymbol{E}_0\otimes \boldsymbol{N}^{-l}$, the curvature form ϕ for $\boldsymbol{N}$ is given by $\psi=\omega-l\phi$ or $\phi=\frac{\sqrt{-1}}{2}\sum\left(\frac{c_i-1}{l}\right)\omega_i \wedge \overline{\omega}_i$. Thus the curvature form ψ_t for $\boldsymbol{E}_t=\boldsymbol{E}_0\otimes \boldsymbol{N}^{-t}$ is given by $\psi_t=\omega-t\phi=\frac{\sqrt{-1}}{2}\sum\left(\frac{t}{l}+\left(\frac{l-t}{l}\right)c_i\right)\omega_i \wedge \overline{\omega}_i$. Consequently, for $0<t\leqq l$, $\psi_t>0$ and so $\boldsymbol{E}_t$ is positive.

We consider now the canonical bundle $\boldsymbol{K}$ of X. By using the bundle along the fibres, we see that $\boldsymbol{K}=\boldsymbol{K}_Y\otimes \boldsymbol{N}^r$ ($\boldsymbol{K}_Y=$line bundle on Y). We want $\boldsymbol{E}_t\otimes \boldsymbol{K}^{-1}$ to be positive for $0<t\leqq l$; by replacing $\boldsymbol{G}$ by $\boldsymbol{G}\otimes \boldsymbol{K}_Y^{-1}$, it will suffice to have $\boldsymbol{E}_t\otimes \boldsymbol{N}^{-r}>0$ for $0<t\leqq l$.

Now we may assume that $0\leqq c_i \leqq 1$, because replacing $\boldsymbol{F}$ by $\boldsymbol{F}^2$ changes ω_i into $\omega_i=\sqrt{2}\omega_i$ and c_i into $c_i'=\frac{c_i}{2}$ The curvature form ω_t for $\boldsymbol{E}_t\otimes \boldsymbol{N}^{-r}$ is

$$\omega_t=\frac{\sqrt{-1}}{2}\sum\left\{\left[\frac{t}{l}+\left(\frac{l-t}{l}\right)c_i\right]-\left[\frac{r}{l}\,(c_i-1)\right]\right\}\omega_i \wedge \overline{\omega}_i.$$

Thus

$$\omega_t \geqq \frac{\sqrt{-1}}{2}\sum\left[\frac{t}{l}+\left(\frac{l-t}{l}\right)c_i\right]\omega_i \wedge \overline{\omega}_i>0 \text{ for } 0<t\leqq l.$$

It follows that $H^1(X, \Omega(\boldsymbol{E}_t))=0$ for $0<t\leqq l$. For $t=0$, we may assume that $H^1(X, \Omega(\boldsymbol{E}_0))=H^1(Y, \Omega(\boldsymbol{G}))=0$. Q.E.D.

(e) Set $\boldsymbol{F}_t=\boldsymbol{F}\otimes \boldsymbol{L}^{l-t}$ so that $\boldsymbol{F}_t \mid X=\boldsymbol{E}_t$.

For $0\leqq t<l$ we have the exact sheaf sequences

$$0\rightarrow\Omega(\boldsymbol{F}_{t+1})\xrightarrow{\sigma}\Omega(\boldsymbol{F}_t)\xrightarrow{\rho}\Omega(\boldsymbol{E}_t)\rightarrow 0 \tag{5.2}$$

where σ is multiplication by the local equation of $X\subset W$. By the lemma in (d) we get $H^1(W, \Omega(\boldsymbol{F}_{t+1}))\rightarrow H^1(W, \Omega(\boldsymbol{F}_t))\rightarrow 0$ for $0\leqq t<l$. But for $t=l-1, \boldsymbol{F}_{t+1}=\boldsymbol{F}$ and so $H^1(W, \Omega(\boldsymbol{F}_l))=0$. It follows that $H^1(W, \Omega(\boldsymbol{F}_t))=0$ for $0\leqq t<l$.

From the exact cohomology sequences of (5.2), we have then the array:

$$(5.3)\qquad \left\{\begin{array}{l} 0\to H^0(\Omega(\boldsymbol{F}_1))\xrightarrow{\sigma} H^0(\Omega(\boldsymbol{F}_0))\xrightarrow{\rho} H^0(\Omega(\boldsymbol{E}_0))\to 0\\ 0\to H^0(\Omega(\boldsymbol{F}_2))\xrightarrow{\sigma} H^0(\Omega(\boldsymbol{F}_1))\xrightarrow{\rho} H^0(\Omega(\boldsymbol{E}_1))\to 0\\ \qquad\qquad\qquad\qquad\vdots\\ \qquad\qquad\qquad\qquad\vdots\\ 0\to H^0(\Omega(\boldsymbol{F}_l))\xrightarrow{\sigma} H^0(\Omega(\boldsymbol{F}_{l-1}))\xrightarrow{\rho} H^0(\Omega(\boldsymbol{E}_{l-1}))\to 0. \end{array}\right.$$

If now $\phi_0,\cdots,\phi_s$ is a basis for $H^0(\Omega(\boldsymbol{E}))$; $\psi_0,\cdots,\psi_t$ is a basis for $H^0(\Omega(\boldsymbol{E}_1))$;$\cdots$;$\gamma_0,\cdots\gamma_\mu$ a basis for $H^0(\Omega(\boldsymbol{F}))$ ($\boldsymbol{F}_l=\boldsymbol{F}$ and $\boldsymbol{E}_0=\boldsymbol{E}$), then

$$(*)\qquad \phi_0,\cdots,\phi_s;\sigma\psi_1,\cdots,\sigma\psi_t;\cdots;\sigma^l\gamma_0,\cdots,\sigma^l\gamma_u$$

will be part of a basis for $H^0(\Omega(\boldsymbol{F}_0))$. Taking the elements in (*) as homogeneous coordinates we get a mapping $\Phi\colon W\to P_N$. Since $\sigma=0$ on X, $\Phi\mid X=\Lambda$; on $W-X$, σ is non-zero so there Φ is given by

$$[\sigma^{-l}\phi_0,\cdots,\sigma^{-l}\phi_s;\cdots;\cdots;\gamma_0,\cdots,\gamma_l].$$

Thus Φ is biregular on $W-X$.

(f) Set $Z=\Phi(W)\subset P_N$. Since $Y\cong\Phi(X)$, there is an embedding $Y\subset Z$ such that: (i) $\Phi(X)=Y$; (ii) $\Phi\mid X=\pi$; and (iii) $\Phi\colon W-X\cong Z-Y$. We want to show that Z is non-singular and that W is the monoidal transform of Z along Y. In order to do this, we assume that $\boldsymbol{G}$ is such that the following holds: For any fibre $P_y\subset X$, if $\mathfrak{J}_y$ is the ideal sheaf of P_y, then, for $q>0$, $H^q(Y,R_\pi{}^0(\Omega(\boldsymbol{N}^{-1})\otimes\mathfrak{J}_y)\otimes\Omega(\boldsymbol{G}))=0$. Indeed, if our original $\boldsymbol{E}\to Y$ does not work, then we may, as before, pass to a suitable power.

Lemma 5.5. *With the above notations, we have that* $H^1(X,\mathfrak{J}_y\cdot\Omega(\boldsymbol{E}_1))=0$ *for all* $y\in Y$.

Proof. If we can show that $R_\pi{}^q(\mathfrak{J}_y\cdot\Omega(\boldsymbol{E}_1))=0$ for $q>0$, then, by the Leray spectral sequence, we have

$$H^1(X,\mathfrak{J}_y\cdot\Omega(\boldsymbol{E}_1))\cong H^1(Y,R_\pi{}^0(\mathfrak{J}_y\cdot\Omega(\boldsymbol{E}_1))).$$

But $\boldsymbol{E}_1=\boldsymbol{E}_0\otimes\boldsymbol{N}^{-1}=\pi^{-1}(\boldsymbol{G})\otimes\boldsymbol{N}^{-1}$ and thus

$$R_\pi{}^0(\mathfrak{J}_y\cdot\Omega(\boldsymbol{E}_1))\cong R_\pi{}^0(\mathfrak{J}_y\cdot\Omega(\boldsymbol{N}^{-1}))\otimes\Omega(\boldsymbol{G}).$$

Because of $H^q(Y,R_\pi{}^0(\Omega(\boldsymbol{N}^{-1})\otimes\mathfrak{J}_y)\otimes\Omega(\boldsymbol{G}))=0$ for $q>0$, we are done.

Since $\boldsymbol{E}_1 \mid \boldsymbol{P}_{y'} = \boldsymbol{H}$, we clearly have $R_\pi{}^q(\mathfrak{J}_y \cdot \Omega(\boldsymbol{E}_1))_{y'} = 0$ for $q > 0$ and $y' \neq y$. Furthermore, $R_\pi{}^q(\mathfrak{J}_y \cdot \Omega(\boldsymbol{E}_1))_y \cong \boldsymbol{m}_y \hat{\otimes} H^q(P_y, \Omega(\boldsymbol{H}))$ for any q and where $\boldsymbol{m}_y \subset \mathfrak{O}_{Y,y}$ is the maximal ideal. This proves the lemma.

(g) Now we can complete the proof. Let $P \subset X$ be a general fibre of $X \xrightarrow{\pi} Y$ and $\mathfrak{J} \subset \mathfrak{O}_X$ the ideal sheaf of P. We have the sequence

$$0 \to \mathfrak{J} \cdot \Omega(\boldsymbol{E}_1) \to \Omega(\boldsymbol{E}_1) \to \Omega_P(\boldsymbol{H}) \to 0$$

and, by the Lemma in (f) we get

$$\begin{tikzcd}
H^0(W, \Omega(\boldsymbol{F}_1)) \arrow[r, "\rho"] \arrow[dr, "\tilde{\rho}"] & H^0(X, \Omega(\boldsymbol{E}_1)) \arrow[r] \arrow[d] & 0 \\
& H^0(P, \Omega(\boldsymbol{H}_1)) \arrow[d] & \\
& 0. &
\end{tikzcd}$$

Let $y_0 \in Y \subset Z$ be a fixed point and set $P = \pi^{-1}(y_0)$. Denote by $U(y_0) \subset W$ a small tubular neighborhood around P. Then we may choose elements $\xi_0, \cdots, \xi_n; \eta_0, \cdots, \eta_r$ from (*) such that: (i) $\xi_0, \cdots, \xi_n \in H^0(W, \Omega(\boldsymbol{F}_0))$ and $\rho(\xi_0), \cdots, \rho(\xi_n) \in H^0(X, \Omega(\boldsymbol{E}_0)) \cong H^0(Y, \Omega(\boldsymbol{G}))$ embed $\pi(X \cap \mathrm{U}(y_0))$ (ii) $\eta_1, \cdots, \eta_r \in H^1(W, \Omega(\boldsymbol{F}_1))$ and $\tilde{\rho}(\eta_1), \cdots, \tilde{\rho}(\eta_r) \in H^0(P_y, \Omega(\boldsymbol{H}))$ give a in a projective space. Also, we may assume that $\xi_0(w) \neq 0$ for $w \in U(y_0)$; basis for $y \in \pi(U(y_0) \cap X)$. Then, by rearranging the elements in (*), we may assume that $\Phi \mid U(y_0)$ has the form

$$\Phi(w) = [\xi_0(w), \cdots, \xi_n(w); \sigma\eta_1(w), \cdots, \sigma\eta_r(w); \cdots].$$

We set $\zeta_i(w) = \xi_i(w)/\xi_0(w)$ $(1 \leqq i \leqq n)$ and $\lambda_\alpha(w) = \dfrac{\sigma\eta_\alpha(w)}{\xi_0(w)}$ $(1 \leqq \alpha \leqq r)$. Then the ζ_i and λ_α are holomorphic functions in $U(y_0)$ and $\lambda_\alpha \mid X = 0$.

Let $V \subset \boldsymbol{C}^{n+r}$ be a polycylinder with coordinates $(\zeta_1, \cdots, \zeta_n; \lambda_1, \cdots, \lambda_r)$, and let P_{r-1} have homogeneous coordinates $[t_1, \cdots, t_r]$. Define $Q(V, P_{r-1}) \subset V \times P_{r-1}$ by

$$Q(V, P_{r-1}) = \{(\zeta, \lambda, t) : \lambda_\alpha t_\beta - \lambda_\beta t_\alpha = 0; 1 \leqq \alpha, \beta \leqq r\}.$$

Then define $F \colon U(y_0) \to Q(V, P_{r-1})$ by

$$F(w) = (\zeta_1(w), \cdots, \zeta_n(w); \lambda_1(w), \cdots, \lambda_r(w); [\eta_1(w), \cdots, \eta_r(w)]).$$

It is easy to check, using Hartog's theorem, that F is biregular on some tubular neighborhood $U \subset U(y_0)$ of $\pi^{-1}(y_0)$. Q. E. D.

3. Reduction to the codimension one case. Let $X \subset W$ be a germ of embedding of a compact, complex manifold X in a complex manifold W. Denote by $\tilde{W}$ the monoidal transform of W along X, let $\tilde{\omega}\colon \tilde{W} \to W$ be the blowing down map, and set $\tilde{X} = \tilde{\omega}^{-1}(X)$. Finally let $\mathfrak{J} \subset \mathfrak{O}_W$ be the ideal sheaf of X and $\tilde{\mathfrak{J}} \subset \mathfrak{O}_{\tilde{W}}$ the ideal sheaf of $\tilde{X}$. The same spectral sequence argument as used in 5.2 just above gives:

PROPOSITION 5.1. *Let $\boldsymbol{E} \to W$ be a holomorphic vector bundle and $\tilde{\boldsymbol{E}} = \tilde{\omega}^{-1}(\boldsymbol{E})$. Then, if $H^q(\tilde{W}, \tilde{\mathfrak{J}}^{\mu} \cdot \Omega(\tilde{\boldsymbol{E}})) = 0$ for $0 \leqq q < n$ and $\mu \geqq \mu_0$, it follows that $H^q(W, \mathfrak{J}^{\mu} \cdot \Omega(\boldsymbol{E})) = 0$ for $0 \leqq q < n$ and $\mu \geqq \mu_0$.*

Remark. In case the normal bundle $\boldsymbol{N}$ of $X \subset W$ is positive, this Proposition says that: (i) we can work on W with positive bundles and non-locally free sheaves; or (ii) we can lift to $\tilde{W}$ and work with locally free sheaves but with bundles of mixed type.

VI. Differential-Geometric Foundations.

1. Local Hermitian geometry. Let X be a complex manifold on which we have an Hermitian metric. Locally we may choose a family $\omega^1, \cdots, \omega^n$ of $(1,0)$ forms such that $ds^2 = \sum \omega^i \bar{\omega}^i$. The *associated 2-form* is then $\gamma = \frac{\sqrt{-1}}{2} \sum \omega^i \wedge \bar{\omega}^i$.

PROPOSITION 6.1. *There exist (locally) a unique set of 1-forms ϕ_j^i which satisfy*

$$\phi_j^i + \bar{\phi}_i^j = 0 \tag{6.1}$$

$$d\omega^i = \sum \phi_j^i \wedge \omega^j + \tfrac{1}{2} \sum T_{jk}^i \omega^j \wedge \omega^k \qquad (T_{jk}^i + T_{kj}^i = 0). \tag{6.2}$$

Proof. We choose local holomorphic coordinates $z^1, \cdots, z^n$ and write $\omega^i = \sum a_j^i dz^j$. In matrix form, this becomes $\omega = a dz$ where $\omega = {}^t(\omega^1, \cdots, \omega^n)$ and $dz = {}^t(dz^1, \cdots, dz^n)$. If we define h by $h_{ij} = (dz^i, dz^j)$, then ${}^t\bar{h} = h$ and $ah{}^t\bar{a} = I$.

Now let $\phi = a(\partial h \cdot h^{-1})a^{-1} + da \cdot a^{-1}$. Then

$$d\omega = da \wedge dz = \phi \wedge \omega - a(\partial h \cdot h^{-1})a^{-1} \wedge \omega. \tag{6.3}$$

Because of (6.3), it will suffice to show that $\phi + {}^t\bar{\phi} = 0$. Given $x_0 \in X$, we may assume that our coordinate system has been chosen so that $h(x_0) = I = a(x_0)$. Then at x_0,

$$\phi + {}^t\bar{\phi} = \partial h + da + \bar{\partial}{}^t\bar{h} + d{}^t\bar{a} = da + dh + d{}^t\bar{a} = d(ah{}^t\bar{a})(x_0) = 0.$$

It is easy to verify that ϕ is unique. Q.E.D.

We write locally

$$\phi_j{}^i = \sum \Lambda_{jk}{}^i \omega^k + \sum \Lambda_{j\bar{k}}{}^i \overline{\omega}^k \tag{6.4}$$

where, by (6.1),

$$\Lambda_{j\bar{k}}{}^i + \bar{\Lambda}_{ik}{}^j = 0. \tag{6.5}$$

Now

$$d\gamma = \frac{\sqrt{-1}}{2}\, d({}^t\omega \wedge \overline{\omega})$$

$$= \frac{\sqrt{-1}}{2} \{ {}^t(\phi \wedge \omega + T) \wedge \overline{\omega} - {}^t\omega \wedge (\overline{\phi \wedge \omega + T}) \} = \sqrt{-1}\,\mathrm{Re}({}^tT \wedge \overline{\omega})$$

and do $d\gamma = 0$ if, and only if, $T = 0$. This T is the so-called torsion tensor which vanishes if, and only if, the metric is Kähler.

We shall write a $(0, q)$ form Ψ on X as $\Psi = \frac{1}{q!} \sum_I \psi_{\bar{I}} \overline{\omega}^I$ where $\bar{I} = (\bar{i}_1, \cdots, \bar{i}_q)$ and $\overline{\omega}^I = \overline{\omega}^{i_1} \wedge \cdots \wedge \overline{\omega}^{i_q}$. The covariant derivatives $\psi_{\bar{I};j}$ and $\psi_{\bar{I};\bar{j}}$ are then defined by

$$\sum \psi_{\bar{I};j} \omega^j + \sum \psi_{\bar{I};\bar{j}} \overline{\omega}^j = d\psi_{\bar{I}} - q \sum \bar{\phi}_{i_1}{}^j \psi_{\bar{j}\bar{i}_2 \cdots \bar{i}_q}. \tag{6.6}$$

The operator $\bar{\partial}: A^q \to A^{q+1}$ is defined as usual and satisfies $\bar{\partial}^2 = 0$. We also define $\bar{\delta}: A^q \to A^{q-1}$ by

$$\int_X \langle \bar{\delta}\psi, \eta \rangle = \int_X \langle \psi, \bar{\partial}\eta \rangle \qquad \text{(support } \eta \text{ compact)} \tag{6.7}$$

Then $\bar{\delta}^2 = 0$ and we have the Laplacian $\Box$ is given by

$$\Box = \bar{\partial}\bar{\delta} + \bar{\delta}\bar{\partial}. \tag{6.8}$$

To obtain a local expression for $\Box$, we define $T: A^q \to A^{q+1}$ by

$$T(\psi) = \frac{1}{(q-1)!} \sum \bar{T}_{jk}{}^i \psi_{\bar{i}\,\bar{i}_2 \cdots \bar{i}_q} \overline{\omega}^j \wedge \overline{\omega}^k \wedge \overline{\omega}^{i2} \wedge \cdots \wedge \overline{\omega}^{iq} \tag{6.9}$$

and $T^{\#}: A^q \to A^{q-1}$ by

$$\langle T^{\#}\psi, \eta \rangle = \langle \psi, T\eta \rangle \qquad (\psi \in A^q, \eta \in A^{q-1}). \tag{6.10}$$

Finally, we let $\Phi_j{}^i = d\phi_j{}^i - \sum \phi_k{}^i \wedge \phi_j{}^k$ be curvature (which is of type (1,1), $\Phi_{i\bar{j}} = \sum \Phi_{ik\bar{j}}{}^k$ the Ricci tensor, and define $\Phi: A^q \to A^q$ by

$$\Phi(\psi) = \frac{1}{(q-1)!} \sum \Phi_{i\bar{j}} \psi_{\bar{i}\,\bar{i}_2 \cdots \bar{i}_q} \overline{\omega}^j \wedge \overline{\omega}^{i2} \wedge \cdots \wedge \overline{\omega}^{iq}. \tag{6.11}$$

Then a calculation gives

$$\text{(6.12)} \qquad \square\psi = \frac{1}{q!}\sum \psi_{\bar{I};\bar{j};j}\bar{\omega}^{I} + [T,\bar{\delta}](\psi) + [T^{\#},\bar{\partial}](\psi) \\ - [T,T^{\#}](\psi) + \Phi(\psi)$$

where $[A,B](\psi) = A(B(\psi)) + B(A(\psi))$. (For example, it follows from (6.12) that $\square = \frac{1}{2}\Delta(\Delta = d\delta + \delta d)$ if, and only if, $T = 0$. This is because $[T,T^{\#}] = 0$ if, and only if, $T = 0$.)

Now suppose that W' is a Hermitian manifold and that $W \Subset W'$ is a *finite submanifold*. Thus ∂W is a smooth, connected, compact real hypersurface. We may define a real function r in a neighborhood of ∂W such that (i) $|dr| = 1$, and (ii) $r < 0$ defines $W \subset W'$. Then $\partial\bar{\partial}r \mid \boldsymbol{T}(\partial W)$ gives the *E. E. Levi form* L of W in W'. Locally, in a neighborhood of a boundary point, we may choose $\omega^1, \cdots, \omega^n$ such that $\omega^n = \partial r$. By (6.2) and (6.5) we get

$$\text{(6.13)} \qquad -\bar{\partial}\omega^n = \sum \Lambda_{j\bar{k}}{}^{n}\omega^j \wedge \bar{\omega}^k$$

so that the matrix $\{\Lambda_{j\bar{k}}{}^{n}\}_{1 \leqq j,k \leqq n-1}$ gives L. For $\psi \in A^q$, we set

$$\text{(6.14)} \qquad L(\psi,\psi) = (L\psi,\psi) = \frac{1}{(q-1)!}\sum_{j,k=1}^{n-1} \Lambda_{j\bar{k}}{}^{n}\psi_{\bar{j}\bar{i}_2\cdots\bar{i}_q}\bar{\psi}_{\bar{k}\bar{i}_2\cdots\bar{i}_q}.$$

2. Holomorphic vector bundles. Let $\boldsymbol{E} \to X$ be a holomorphic vector bundle and $P \xrightarrow{\pi} X$ the holomorphic principal bundle. A connexion in $\boldsymbol{E}$ is given by a matrix-valued one-form $\theta = \{\theta^{\rho}{}_{\sigma}\}$ on P which satisfies the usual conditions. Given θ, we write $\theta = \theta' + \theta''$ where $\theta'(\theta'')$ is the $(1,0)$ (resp. $(0,1)$) part of θ; we say that θ defines a *complex connexion* if $\theta'' = 0$. As usual, a connexion θ induces a covariant differential

$$D\colon A^{p,q}(\boldsymbol{E}) \to A^{p+1,q}(\boldsymbol{E}) \oplus A^{p,q+1}(\boldsymbol{E})$$

and we write $D = D' + D''$ for the type decomposition. A complex connexion is one for which $D'' = \bar{\partial}$.

Suppose now that we have in $\boldsymbol{E} \to X$ an Hermitian metric. The following is known:

PROPOSITION 6.2. *There exists a unique complex connexion θ which is compatible with the metric in $\boldsymbol{E}$. Furthermore, the curvature Θ of θ is of type* $(1,1)$.

Proof. P is the manifold of holomorphic $\boldsymbol{E}$-frames $(e_1, \cdots, e_r)$ and we define on P an Hermitian matrix h by $h_{\rho\sigma} = (e_\rho, e_\sigma)$. Now D and θ are related by $De_\rho = \sum \theta_\rho{}^\sigma e_\sigma$ and to say that θ is compatible with the metric means that $d(e,e') = (De,e') + (e,De')$ for sections e, e' of $\boldsymbol{E}$. Thus θ is a solution of the system

$$\begin{cases} dh_{\rho\sigma} = \sum \theta_\rho{}^\tau h_{\tau\sigma} + \sum h_{\rho\tau}\bar{\theta}_\sigma{}^\tau \\ \theta'' = 0. \end{cases} \tag{6.15}$$

The system (6.15) has the unique solution

$$\theta_\sigma{}^\rho = \sum \partial h_{\sigma\tau}(h^{-1})_{\tau\rho}. \tag{6.16}$$

Remarks. (i) If $P^\# \subset P$ is the real sub-bundle of unitary frames, $\theta \mid P^\#$ is given by one-forms $\theta_\sigma{}^\rho$ which satisfy $\theta_\sigma{}^\rho + \bar{\theta}_\rho{}^\sigma = 0$; then

$$\Theta_\sigma{}^\rho = d\theta_\sigma{}^\rho - \sum \theta_\tau{}^\rho \wedge \theta_\sigma{}^\tau;$$

and

$$\Theta_\sigma{}^\rho + \bar{\Theta}_\rho{}^\sigma = 0 \tag{6.17}$$

(ii) To interpret the second statement in Proposition 6.2, we suppose that $\boldsymbol{E} \to X$ is a *differentiable* bundle with fibre $\boldsymbol{C}^r$ and that θ on P gives a connexion for $\boldsymbol{E}$. If $\omega^1, \cdots, \omega^n$ are a local $(1,0)$ co-frame on X, then the differential system generated by the $\theta_\sigma{}^\rho$ and ω^i gives an almost-complex structure J^θ on P.

PROPOSITION. *If the curvature Θ of θ is of type $(1,1)$, then J^θ is integrable.*

Proof. Since $d\theta_\sigma{}^\rho - \sum \theta_\tau{}^\rho \wedge \theta_\sigma{}^\tau = \Theta_\sigma{}^\rho$, we see that, if Θ is of type $(1,1)$, then $d\theta_\sigma{}^\rho$ is in the exterior ideal generated by $\{\theta_\sigma{}^\rho; \omega^i\}$.

Example. Suppose that X is a compact Kähler manifold and that $\gamma \in H^2(X,\boldsymbol{Z})$ is a cohomology class of type $(1,1)$. Then there exists a differentiable line bundle (with fibre $\boldsymbol{C}$) $\boldsymbol{E} \to X$ such that $c_1(\boldsymbol{E}) = \gamma$. Let γ be a real form of type $(1,1)$ which represents γ and let θ be a connexion for $\boldsymbol{E}$. Then $\dfrac{\sqrt{-1}}{2\pi}\, d\theta = \gamma + d\eta$ where η is a global one-form on X. Thus $\theta - \dfrac{2\pi}{\sqrt{-1}}\eta$ is a connexion in $\boldsymbol{E}$ whose curvature γ is of type $(1,1)$ and so $\boldsymbol{E}$ is a holomorphic bundle. This is a theorem of Kodaira-Spencer [14].

(iii) *Example.* Let $\boldsymbol{E} \xrightarrow{\pi} X$ be an Hermitian holomorphic vector bundle and suppose that we have on X an Hermitian metric ds^2. Then $\boldsymbol{E}^\# = \pi^{-1}(\boldsymbol{E})$ is a holomorphic bundle over $\boldsymbol{E}$ and there is a *canonical section* $\sigma \in H^0(\boldsymbol{E}, \Omega(\boldsymbol{E}^\#))$ defined by $\sigma(z,\xi) = (z,\xi;\xi)$ $(z \in X, \xi \in \boldsymbol{E}_z)$. We define on $\boldsymbol{E}$ an Hermitian metric $\phi = \pi^*(ds^2) + (D\sigma, D\sigma)$ where D is the covariant differential associated to the metric connexion θ.

PROPOSITION. *ϕ is a Kähler metric if, and only if, ds^2 is Kähler and the curvature Θ of θ is zero.*

Proof. Assume ds^2 is Kähler and denote by $\langle D\sigma, D\sigma\rangle$ the exterior form associated to $(D\sigma, D\sigma)$. Then

$$d\langle D\sigma, D\sigma\rangle = \langle D^2\sigma, D\sigma\rangle + \langle D\sigma, D^2\sigma\rangle = 2Re\langle \Theta\sigma, D\sigma\rangle = 0$$

if, and only if, $\Theta = 0$. Q. E. D.

Suppose that $\boldsymbol{E} \to X$ is a holomorphic Hermitian vector bundle over the Hermitian manifold X, and let $e_1, \cdots, e_r; \omega^1, \cdots, \omega^n$ be local unitary frames for $\boldsymbol{E}, \boldsymbol{T}^*$ respectively. Then $\phi \in A^q(\boldsymbol{E})$ is written as $\phi = \frac{1}{q!}\sum \phi_{\bar{I}}{}^{\rho}\bar{\omega}^I \otimes e_\rho$. If $\theta = (\theta_\sigma{}^\rho)$ is the connexion form of the metric connexion, then the covariant derivatives of ϕ are given by

$$\phi_{\bar{I}|j}{}^{\rho} = \phi_{\bar{I};j}{}^{\rho} + \sum \theta_{\sigma j}{}^{\rho}\phi_{\bar{I}}{}^{\sigma} \qquad (\theta_\sigma{}^\rho = \sum \theta_{\sigma i}{}^{\rho}\omega^i + \sum \theta_{\sigma\bar{j}}{}^{\rho}\bar{\omega}^j); \tag{6.18}$$

$$\phi_{\bar{I}|\bar{j}}{}^{\rho} = \phi_{\bar{I};\bar{j}}{}^{\rho}. \tag{6.19}$$

The operator $\vartheta \colon A^q(\boldsymbol{E}) \to A^{q-1}(\boldsymbol{E})$ is defined by

$$(\vartheta\phi)^\rho = \frac{(-1)^q}{(q-1)!}\sum \phi_{\bar{i}\cdots\bar{i}_{q-1}\bar{j}|j}{}^{\rho}\bar{\omega}^{i_1} \wedge \cdots \wedge \bar{\omega}^{i_{q-1}} - T^{\#}(\phi). \tag{6.20}$$

Then $\vartheta^2 = 0$ and the Laplacian $\square = \bar{\partial}\vartheta + \vartheta\bar{\partial}$ is given as

$$(\square\phi)^\rho = \frac{1}{q!}\sum \phi_{\bar{I}|\bar{j}|j}{}^{\rho}\bar{\omega}^I + Q(\nabla\phi) + R(\bar{\nabla}\phi) + S(\phi) + \Theta(\phi) \tag{6.21}$$

where

$$Q(\nabla\phi) = [T, \vartheta](\phi);\ R(\bar{\nabla}\phi) = [T^{\#}, \bar{\partial}](\phi);$$
$$S(\phi) = -[T, T^{\#}](\phi) + \Phi(\phi);$$

and

$$\Theta(\phi)^\rho = \frac{1}{(q-1)!}\sum \Theta_{\sigma i\bar{j}}{}^{\rho}\phi_{\bar{i}\,\bar{i}_2\cdots\bar{i}_q}{}^{\sigma}\bar{\omega}^j \wedge \bar{\omega}^{i_2} \wedge \cdots \wedge \bar{\omega}^{i_q}. \tag{6.22}$$

3. Induced connexions in sub- and -quotient bundles. Let $\boldsymbol{E} \to X$ be a holomorphic Hermitian vector bundle and suppose given an exact sequence of holomorphic bundles

$$0 \to \boldsymbol{S} \to \boldsymbol{E} \to \boldsymbol{Q} \to 0. \tag{6.23}$$

Let r = fibre dimension of $\boldsymbol{E}$, s = fibre dimension of $\boldsymbol{S}$, and we agree an the ranges of indices $1 \leqq \rho, \sigma \leqq r; 1 \leqq \alpha, \beta \leqq s; s+1 \leqq i, j \leqq r$. Also we denote by $B \subset P^{\#}$ the unitary frames $(e_1, \cdots, e_r)$ for $\boldsymbol{E}$ such that $e_1, \cdots, e_s$ lie in $\boldsymbol{S}$. Then the metric connexion $\theta = \{\theta^{\rho}\sigma\}$ when restricted to B, defines connexions $\theta_S = \{\theta_\beta{}^\alpha\}$ and $\theta_Q = \{\theta_j{}^i\}$ in $\boldsymbol{S}$ and $\boldsymbol{Q}$ respectively. Thus, for $\xi = \sum \xi^\alpha e_\alpha \in A^0(\boldsymbol{S})$, $\eta = \sum \eta^i e_i \in A^0(\boldsymbol{Q})$, $D_S \xi = \sum (d\xi^\alpha + \theta_\beta{}^\alpha \xi^\beta) e_\alpha$ and $D_Q \eta = \sum (d\eta^i + \theta_j{}^i \eta^j) e_i$.

LEMMA 6.2. $D - D_S \in A^{1,0}(\mathrm{Hom}(\boldsymbol{S}, \boldsymbol{Q}))$.

Proof. For $\xi \in A^0(\boldsymbol{S})$ as above, $((D - D_S)\xi = \sum \theta_\alpha{}^i \xi^\alpha e_i$ and so $D - D_S$ is a 1-form with values in $\mathrm{Hom}(\boldsymbol{S}, \boldsymbol{Q})$. For $\zeta \in \boldsymbol{S}_{x_0}$, there exists a holomorphic section $\xi \in \Omega(\boldsymbol{S})_{x_0}$ with $\xi(x_0) = \zeta$. Then $(D''\xi)(x_0) = 0$ which gives $0 = (\sum \theta_\alpha{}^{i''} \xi^\alpha e_i)(x_0) = \sum \theta_\alpha{}^{i''} \zeta^\alpha e_i = (D - D_S)''\zeta$ and so $(D - D_S)'' = 0$; i.e. $D - D_S \in A^{1,0}(\mathrm{Hom}(\boldsymbol{S}, \boldsymbol{Q}))$. Q.E.D.

Definition. The operator $b = D - D_S \in A^{1,0}(\mathrm{Hom}(\boldsymbol{S}, \boldsymbol{Q}))$ is defined to be the *second fundamental form* of $\boldsymbol{S}$ in $\boldsymbol{E}$.

Remark. We shall give below two intepretations of b.

From $\theta_\alpha{}^{i''} = 0$ and $\theta_\alpha{}^i + \bar{\bar{\theta}}_i{}^\alpha = 0$, we get $\theta_i{}^{\alpha'} = 0$. Thus

$$c = D - D_Q \in A^{0,1}(\mathrm{Hom}(\boldsymbol{Q}, \boldsymbol{S}))$$

and is given by $c = -{}^t\bar{b}$.

PROPOSITION 6.3. *θ_S gives the metric connexion of $\boldsymbol{S}$ with the induced metric.*

Proof. Since $\theta_\beta{}^\alpha + \bar{\theta}_\alpha{}^\beta = 0$ we need only verify that $D_S'' = \bar{\partial}$. By Lemma 6.2, $D_S'' = D'' = \bar{\partial}$. Q.E.D.

The curvature Θ_S of θ_S is given by

$$\begin{aligned}(\Theta_S)_\beta{}^\alpha = d\theta_\beta{}^\alpha - \sum \theta_\gamma{}^\alpha \wedge \theta_\beta{}^\gamma &= \Theta_\beta{}^\alpha + \sum \theta_i{}^\alpha \wedge \theta_\beta{}^i \\ &= \Theta_\beta{}^\alpha - \sum \theta_i{}^\alpha \wedge \bar{\theta}_i{}^\beta, \text{ i.e.}\end{aligned}$$

$$\Theta_S = \Theta - b \wedge {}^t\bar{b}. \tag{6.24}$$

Similarly,

$$\Theta_Q = \Theta + {}^t b \wedge \bar{b}. \tag{6.25}$$

Examples. (i) Suppose that $\boldsymbol{E}$ is the trivial bundle and take the flat metric connexion. Then, in the notation of § IV.4, there is a holomorphic mapping $f: X \to G(s, r-s)$ such that $f^{-1}(\boldsymbol{E}(s, r-s)) = \boldsymbol{Q}$ and $b \in \mathrm{Hom}(\boldsymbol{T}(X), \mathrm{Hom}(\boldsymbol{S}, \boldsymbol{Q}))$ is the differential f_* of f. In this case, $\Theta = 0$ and

(6.26) $$\Theta_Q = {}^t b \wedge \bar{b}.$$

(ii) The metric in $\boldsymbol{E}$ defines a C^∞ splitting of (6.23) which is given by $\gamma\colon \boldsymbol{Q} \to \boldsymbol{E}$. Then $\bar{\partial}\gamma \in A^{0,1}(\mathrm{Hom}(\boldsymbol{Q}, \boldsymbol{S}))$ and represents in

$$H^1(X, \Omega(\mathrm{Hom}(\boldsymbol{Q}, \boldsymbol{S})))$$

the cohomology extension class defined by (6.23).

Proposition. $c \equiv \bar{\partial}\gamma$.

Proof. If $e_1^*, \cdots, e_r^*$ is dual to $e_1, \cdots, e_r$, then $\gamma = \sum e_i \otimes e_i^*$. Thus

$$\begin{aligned}\bar{\partial}\gamma &= D''\gamma \\ &= \sum \theta_i^{j\prime\prime} e_j \otimes e_i^* + \sum \theta_i^{\alpha\prime\prime} e_\alpha \otimes e_i^* - \sum \theta_k^{i\prime\prime} e_i \otimes e_k^* - \sum \theta_\alpha^{i\prime\prime} e_i \otimes e_\alpha^* \\ &= \sum \theta_i^{\alpha\prime\prime} e_\alpha \otimes e_i^* = \mathrm{c}. \quad \text{Q.E.D.}\end{aligned}$$

(iii) In example (ii) we have given a cohomological interpretation of the second fundamental form of $\boldsymbol{S}$ in $\boldsymbol{E}$, and we shall now give a geometric interpretation which will justify the terminology. Let W be an Hermitian manifold of dimension r and $X \subset W$ a complex submanifold of dimension s. Let $\boldsymbol{T} = \boldsymbol{T}(W) \mid X$ so that, over X, we have

(6.27) $$0 \to \boldsymbol{T}(X) \to \boldsymbol{T} \to \boldsymbol{N} \to 0$$

where $\boldsymbol{N}$ is the normal bundle of X in W. Along X we choose co-frames $\omega^1, \cdots, \omega^r$ for W such that $\omega^{s+1} = 0, \cdots, \omega^r = 0$ on X. Let $\phi = (\phi_\sigma{}^\rho)$ be the Hermitian connexion on W (§ VI.1) which is also the metric connexion in $\boldsymbol{T}(W)$. If $R \subset B$ is the bundle of unitary frames for X, then R is defined by $\omega^i = 0$ $(i = s+1, \cdots, r)$ and, on R, $\phi_\alpha{}^i = \sum A_{\alpha\beta}{}^i \omega^\beta + \sum C_{\alpha\bar{\beta}}{}^i \bar{\omega}^\beta$. But, on R, $0 = d\omega^i = \sum \phi_\alpha{}^i \wedge \omega^\alpha + \sum T_{\alpha\beta}{}^i \omega^\alpha \wedge \omega^\beta$ (by (6.2)) and it follows that $C_{\alpha\bar{\beta}}{}^i \equiv 0$ and $\phi_\alpha{}^i = \sum A_{\alpha\beta}{}^i \omega^\beta$. If the metric is Kähler $(T = 0)$, then we even get $A_{\alpha\beta}{}^i = A_{\beta\alpha}{}^i$. The quadratic form

(6.28) $$\pi = \sum A_{\alpha\beta}{}^i e_i \otimes \omega^\alpha \otimes \omega^\beta$$

has values in $\boldsymbol{N}$ and, by direct analogy with the Riemannian case, may be called the *second fundamental form of X in W*. Clearly π is also the second fundamental form of $\boldsymbol{T}(X)$ in $\boldsymbol{T}$.

4. Comparison of Hermitian forms. Let X be an Hermitian manifold and suppose we are given a real (1,1) form $A = \dfrac{\sqrt{-1}}{2} \sum a_{ij} \omega^i \wedge \bar{\omega}^j$ $(a_{ij} + \bar{a}_{ji} = 0)$. Then A defines an endomorphism A of $\boldsymbol{T}(X)$ by the rule

$(Ae_i, e_j) = a_{ij}$ where $e_1, \cdots, e_n$ is dual to $\omega^1, \cdots, \omega^n$. Since $A = {}^t\bar{A}$, we may choose our frames $e_1, \cdots, e_n$ so that $A = \frac{\sqrt{-1}}{2}\sum \lambda_i\omega^i \wedge \bar{\omega}^i$ where the λ_j are real.

Definition. A has signature (s, t) if, at each point, s of the λ_j are positive and t are negative.

In this case we may assume that

$$(6.29) \qquad A = \frac{\sqrt{-1}}{2}\{\sum_{1\leqq i\leqq s} \lambda_i\omega^i \wedge \bar{\omega}^i - \sum_{n-t<j\leqq n} \lambda_j\omega^j \wedge \bar{\omega}^j\} \qquad (\lambda_k > 0).$$

The point operators $L\colon A^{p,q} \to A^{p+1,q+1}$ and $\Lambda\colon A^{p,q} \to A^{p-1,q-1}$ are given by $L(\eta) = \gamma \wedge \eta$ and $(\Lambda\eta, \psi) = (\eta, L\psi)$ where $\gamma = \frac{\sqrt{-1}}{2}\sum \omega^i \wedge \bar{\omega}^i$. Suppose now that $A = \frac{\sqrt{-1}}{2}\sum a_{ij}\omega^i \wedge \bar{\omega}^j$ is a real $(1,1)$ form and define the following pointwise Hermitian forms:

$$(6.30) \qquad A(\eta, \eta) = \langle A\eta, \eta\rangle = \frac{1}{(q-1)!}\sum a_{ij}\eta_{i\bar{k}_2\cdots\bar{k}_q}\bar{\eta}_{j\bar{k}\cdots\bar{k}_q}$$

$$(6.31) \qquad A^{\#}(\eta, \eta) = \langle \Lambda A \wedge \eta, \eta\rangle.$$

LEMMA 6.3.

$$(6.32) \qquad A^{\#}(\eta, \eta) = (\operatorname{Trace} A)\,|\,\eta\,|^2 - A(\eta, \eta) = A(*\bar{\eta}, *\bar{\eta}).$$

Proof. Write $A = \frac{\sqrt{-1}}{2}\sum_{m=1}^{n} \lambda_m\omega^m \wedge \bar{\omega}^m$. Then, for $L \subset [1, \cdots, n]$, $A \wedge \bar{\omega}^L = \frac{\sqrt{-1}}{2}\sum_{m\notin L}\lambda_m\omega^m \wedge \bar{\omega}^m \wedge \bar{\omega}^L$ and $\Lambda A \wedge \bar{\omega}^L = (\sum_{m\notin L}\lambda_m)\bar{\omega}^L$. Now

$$\langle \Lambda A \wedge \bar{\omega}^L, \bar{\omega}^{L'}\rangle = 0 \text{ for } L \neq L'$$

and

$$\begin{aligned}\langle \Lambda A \wedge \bar{\omega}^L, \bar{\omega}^L\rangle \\ &= (\sum_{m\notin L}\lambda_m)\,|\,\bar{\omega}^L\,|^2 = (\sum_{m=1}^{n}\lambda_m)\,|\,\bar{\omega}^L\,|^2 - (\sum_{m\in L}\lambda_m)\,|\,\bar{\omega}^L\,|^2 \\ &= (\operatorname{Trace} A)\,|\,\bar{\omega}^L\,|^2 - A(\bar{\omega}^L, \bar{\omega}^L).\end{aligned}$$

Since $A(\bar{\omega}^L, \bar{\omega}^{L'}) = 0$ for $L \neq L'$, this proves the first part of (6.32). The second part may be checked directly.

PROPOSITION 6.4. *Assume that A has signature (s, t). Then there* *exist Hermitian metrics on X such that $A(\eta,\eta)$ is negative definite on A^q for $q > n-t$ and $A^{\#}(\eta,\eta)$ is positive definite on A^q for $q < s$.*

Proof. Write $A = \frac{\sqrt{-1}}{2} \sum_{k=1}^{n} \lambda_k \omega^k \wedge \bar{\omega}^k$. Then, for positive real functions $c_1, \cdots, c_n$, the Hermitian form $\sum c_i \omega^i \bar{\omega}^i$ is well-defined provided that $c_i(x) = c_j(x)$ whenever $\lambda_i(x) = \lambda_j(x)$. Thus, given a constant $c > 0$, we may find a new metric $ds^2 = \sum \phi^i \bar{\phi}^i$ on X such that

$$A = \frac{\sqrt{-1}}{2} \{ c \sum_{1 \leqq i \leqq s} \phi^i \wedge \bar{\phi}^i - \frac{1}{c} \sum_{n-t < j \leqq n} \phi^j \wedge \bar{\phi}^j \}. \tag{6.33}$$

We consider increasing sets of indices $I = (i_1, \cdots, i_\alpha)$, $J = (j_1, \cdots, j_\beta)$, $K = (k_1, \cdots, k_\gamma)$ with

$$I \subset [1, \cdots, s]; \ J \subset [s+1, \cdots, n-t], \ K \subset [n-t+1, \cdots, n].$$

If $\bar{\phi}^{IJK} = \bar{\phi}^I \wedge \bar{\phi}^J \wedge \bar{\phi}^K$, then $A(\bar{\phi}^{IJK}, \bar{\phi}^{I'J'K'}) = 0$ if $IJK \neq I'J'K'$ and

$$A(\bar{\phi}^{IJK}, \bar{\phi}^{IJK}) = (\alpha c - \frac{\gamma}{c}) \,|\, \bar{\phi}^{IJK} \,|^2.$$

If now $\alpha + \beta + \gamma = q$, then $\gamma \geqq q - (n-t)$ so that, if $q > (n-t)$, we may choose $0 < c < q^{\frac{1}{2}}$ and then $\alpha c - \frac{\gamma}{c} < 0$. Thus $A(\eta,\eta)$ is negative definite for $q > n-t$.

If, on the other hand, $q > n-s$ and $c > q^{\frac{1}{2}}$ then $A(\eta,\eta)$ is positive definite on A^q. Thus, if $n-q > n-s$, $A(*\bar{\eta}, *\bar{\eta})$ will be positive definite for $\eta \in A^q$. By (6.32), if $q < s$ and $\eta \in A^q$, $A^{\#}(\eta,\eta) = A(*\bar{\eta}, *\bar{\eta})$ will be positive definite as required. Q. E. D.

Now let $\boldsymbol{E} \to X$ be a holomorphic Hermitian vector bundle. Then $A \in A^{1,1}(\mathrm{Hom}(\boldsymbol{E}, \boldsymbol{E}))$ is said to be Hermitian if $\langle Ae, e' \rangle + \langle e, Ae' \rangle = 0$ for all $e, e' \in \boldsymbol{E}$. Locally we may choose orthonormal frames and write A as $A = \sum A_{\sigma i j}{}^{\rho} e_\rho \otimes e_\sigma^* \otimes \omega^i \wedge \bar{\omega}^j$, and to say that A is Hermitian means that

$$A_{\sigma i j}{}^{\rho} = \bar{A}_{\rho j i}{}^{\sigma} \tag{6.34}$$

By (6.17), the curvature Θ of the metric connexion is Hermitian.

We now define the following Hermitian forms for $\eta \in A^q(\boldsymbol{E})$:

$$A(\eta,\eta) = (A\eta, \eta) = \frac{1}{(q-1)!} \sum A_{\sigma i j}{}^{\rho} \eta_{\bar{i}\bar{k}_1 \cdots k_{q-1}}{}^{\sigma} \overline{\eta_{\bar{j}\bar{k}_1 \cdots \bar{k}_{q-1}}{}^{\rho}} \tag{6.35}$$

$$A^{\#}(\eta,\eta) = (\Lambda A \wedge \eta, \eta) = \frac{1}{q!} \sum \langle \Lambda(A_{\sigma i j}{}^{\rho} \omega^i \wedge \omega^j \wedge \eta), \eta \rangle. \tag{6.36}$$

Observe that (6.36) is compatible with (6.22) and (6.14).

For $e \in \boldsymbol{E}$, $A\{e\} = \frac{\sqrt{-1}}{2} \langle Ae, e \rangle$ is a real $(1,1)$ form and we say that A has signature (s,t) if, for every $e \neq 0$, $A\{e\}$ has signature (s,t). The geometric meaning of this will be given in Proposition 6.5 below.

In the same way as we proved Lemma 6.3, we can show: For $\eta \in A^q(\boldsymbol{E})$,

$$A^{\#}(\eta,\eta) = (\mathrm{Trace}\, A)(\eta,\eta) - A(\eta,\eta) = A(\#\eta, \#\eta) \tag{6.37}$$

where $(\mathrm{Trace}\, A)(\eta,\eta) = \frac{1}{q!} \sum A_{\sigma}{}^{\rho} \eta_{I}{}^{\sigma} \bar{\eta}_{I\rho}$ and $\#\colon A^q(\boldsymbol{E}) \to A^{n-q}(\boldsymbol{E}^*)$ is the usual anti-isomorphism.

5. Positive and negative bundles. Let X be a complex Hermitian manifold and $\boldsymbol{E} \to X$ a holomorphic Hermitian vector bundle. The curvature $\Theta = \{\sum \Theta_{\sigma i j}{}^{\rho} \omega^i \wedge \bar{\omega}^j\} \in A^{1,1}(\mathrm{Hom}(\boldsymbol{E},\boldsymbol{E}))$ is an Hermitian operator and $\boldsymbol{E}$ is defined to be *weakly positive* if Θ has signature $(n,0)$. We say that $\boldsymbol{E}$ is *weakly negative* if $\boldsymbol{E}^*$ is weakly positive; this is the same as saying that Θ has signature $(0,n)$. More generally, we will say that $\boldsymbol{E}$ *has signature* (s,t) if Θ has signature (s,t).

Examples and Remarks. (i) $\boldsymbol{E}$ has signature (s,t) if, and only if, $\boldsymbol{E}^*$ has signature (t,s).

(ii) The condition that $\boldsymbol{E}$ be weakly positive is that the form

$$\Theta(\xi \otimes \eta, \xi \otimes \eta) = \sum \Theta_{\sigma i j}{}^{\rho} \xi^{\sigma} \bar{\xi}_{\rho} \eta^{i} \bar{\eta}^{j} \tag{6.38}$$

on *split tensors* $\xi \otimes \eta$ should be positive definite. We say that $\boldsymbol{E} \to X$ is positive if the forms $\Theta^{\#}(\eta,\eta)$ $(\eta \in A^q(\boldsymbol{E}), 0 \leqq q \leqq n-1)$ are positive definite on all tensors. In [8] it is proven that, if $\boldsymbol{E}$ is weakly positive and if X is compact, then high symetric powers $\boldsymbol{E}^{\mu}$ $(\mu \geqq \mu_0)$ are positive. For example, $\boldsymbol{T}(P_n)$ is weakly positive but *not* positive for $n > 1$.

(iii) Suppose that we have a short exact sequence (6.23). Then the second fundamental form $b \in \mathrm{Hom}(\boldsymbol{T}(X) \otimes \boldsymbol{S}, \boldsymbol{Q})$ and ${}^t b \in \mathrm{Hom}(\boldsymbol{T}(X) \otimes \boldsymbol{Q}, \boldsymbol{S})$ are defined. For $\eta \in \boldsymbol{T}(X)$, $\xi \in \boldsymbol{S}$, $\zeta \in \boldsymbol{Q}$, we have by (6.24), (6.25)

$$\Theta_S(\xi \otimes \eta, \xi \otimes \eta) = \Theta(\xi \otimes \eta, \xi \otimes \eta) - |b(\xi \otimes \eta)|^2 \tag{6.39}$$

$$\Theta_Q(\zeta \otimes \eta, \zeta \otimes \eta) = \Theta(\zeta \otimes \eta, \zeta \otimes \eta) + |{}^t b(\zeta \otimes \eta)|^2. \tag{6.40}$$

Thus a sub-bundle of a weakly negative bundle is weakly negative and a quotient bundle of a weakly positive bundle is weakly positive.

In case $\boldsymbol{E}$ is the trivial bundle and $f\colon X \to G(s, r-s)$ is the associated mapping (§ VI. 3, example 1), it is not true that f_* injective implies that $\boldsymbol{Q}$

is weakly positive (because, e.g., $\boldsymbol{E}(s, r-s) \to G(s, r-s)$ is not weakly positive if $r-s > 1$). It is, however, true that if $\boldsymbol{Q} \to X$ is ample (in the sense of § IV. 4), then it is weakly positive. (*Proof*: Under the isomorphism $\operatorname{Hom}(\boldsymbol{T}(X) \otimes \boldsymbol{S}, \boldsymbol{Q}) \cong \operatorname{Hom}(\boldsymbol{S}, \boldsymbol{Q} \otimes \boldsymbol{T}(X)^*)$, b defines an element $\beta \in \operatorname{Hom}(\boldsymbol{S}, \boldsymbol{Q} \otimes \boldsymbol{T}(X)^*)$ which is, by definition, surjective. Let now $\gamma_1, \cdots, \gamma_s$ be an orthonormal basis for $\boldsymbol{S}_x$ and let $\zeta \in \boldsymbol{Q}_x$, $\eta \in \boldsymbol{T}(X)^*_x$. Then

$$\Theta_{\boldsymbol{Q}}(\zeta \otimes \eta, \zeta \otimes \eta) = \sum_{\alpha=1}^{s} |\langle \beta(\gamma_\alpha), \zeta \otimes \eta \rangle|^2. \tag{6.41}$$

The assertion now follows from (6.41).)

(iv) Suppose that $\boldsymbol{E}$ and $\boldsymbol{F}$ are holomorphic Hermitian vector bundles and let $e_1, \cdots, e_r; f_1, \cdots, f_s$ be bases for $\boldsymbol{E}_x$, $\boldsymbol{F}_x$ $(x \in X)$. If

$$\phi = \sum \phi^{\rho\alpha} e_\rho \otimes f_\alpha \in (\boldsymbol{E} \otimes \boldsymbol{F})_x,$$

we define $\phi_\alpha \in \boldsymbol{E}_x$ by $\phi_\alpha = \sum \phi^{\rho\alpha} e_\rho$ and $\phi_\rho \in \boldsymbol{F}_x$ by $\phi_\rho = \sum \phi^{\rho\alpha} f_\alpha$. Then, for $\eta \in \boldsymbol{T}(X)_x$,

$$\Theta_{\boldsymbol{E}\otimes\boldsymbol{F}}(\phi \otimes \eta, \phi \otimes \eta) = \sum_{\alpha} \Theta_{\boldsymbol{E}}(\phi_\alpha \otimes \eta, \phi_\alpha \otimes \eta) + \sum_{\rho} \Theta_{\boldsymbol{F}}(\phi_\rho \otimes \eta, \phi_\rho \otimes \eta). \tag{6.42}$$

From (6.42) it follows that the tensor product of weakly positive bundles is again weakly positive.

(v) (cf. [3], § 23) Let $\boldsymbol{E} \xrightarrow{\pi} X$ be a holomorphic Hermitian vector bundle and define a real function ϕ on $\boldsymbol{E}$ by $\phi(x, \xi) = |\xi|_x^2 (\xi \in \boldsymbol{E}_x)$. Let $W \subset \boldsymbol{E}$ be defined by $\{(x, \xi) : \phi(x, \xi) < 1\}$. Then, if $\boldsymbol{E}$ has signature (s, t), the E. E. Levi form of ∂W has signature $(t + r - 1, s)$. (*Proof*: Let $x_0 \in X$ be fixed and let $z = (z^1, \cdots, z^n)$ be holomorphic coordinates in a neighborhood U of x_0. If we choose a trivialization $\pi^{-1}(U) \cong U \times \boldsymbol{C}^r$, then the metric is given by an Hermitian matrix $h = (h_{\rho\sigma}(z))$ and we may assume that $h_{\rho\sigma}(o) = \delta_\sigma{}^\rho$, $dh((o)) = 0$. Let $e_0 = (x_0, \xi_0) \in \partial W \cap \boldsymbol{E}_{x_0}$. Since $\phi(z, \xi) = \sum h^\rho{}_\sigma \xi^\rho \bar{\xi}^\sigma$,

$$(\partial\bar{\partial}\phi)(e_0) = \sum d\xi^\rho \wedge d\bar{\xi}^\rho - \sum \Theta_{\sigma i \bar{j}}{}^\rho \xi^\sigma \bar{\xi}^\rho \, dz^i \wedge d\bar{z}^j \tag{6.43}$$

and $\boldsymbol{T}_{e_0}(\partial W)$ is given by the vectors $\gamma = (\zeta^1, \cdots, \zeta^r; \eta^1, \cdots, \eta^n)$ which satisfy

$$\sum \xi_0{}^\rho \bar{\zeta}^\rho = 0 \tag{6.44}$$

thus $\partial\bar{\partial}\phi$ has, on $\boldsymbol{T}_{e_0}(\partial W)$, signature $(t + r - 1, s)$.)

(v) Finally, let $P(\boldsymbol{E}^*) \to X$ be the bundle of projective spaces associated to $\boldsymbol{E}^* \to X$ and $\boldsymbol{L} \to P(\boldsymbol{E}^*)$ the hyperplane bundle (cf. § A. 3). Then the computation in [8], § 9 gives

PROPOSITION 6. 5. *If* $\boldsymbol{E} \to X$ *has signature* (s, t), *then* $\boldsymbol{L} \to P(\boldsymbol{E}^*)$, *with the induced metric, has signature* $(s + r - 1, t)$.

VII. Integral Formulae and Vanishing Theorems for Complex Manifold.

1. Harmonic integrals on open manifolds [15]. Let $W \subset W'$ be a finite complex manifold and $\boldsymbol{E} \to W$ a holomorphic vector bundle. We assume that we have Hermitian metrics in $\boldsymbol{E}$ and on W and that these metrics are *finite* in the sense that they extend to metrics in a neighborhood of W in W'. Denote by $\boldsymbol{A}^q(\boldsymbol{E})$ the space of $\boldsymbol{E}$-valued $(0, q)$ forms which are C^∞ on $\bar{W}$ $(= W \cup \partial W)$. If $\phi, \psi \in \boldsymbol{A}^q(\boldsymbol{E})$, we let

$$(\phi, \psi) = \int_W \langle \phi, \psi \rangle d\mu = \int_W {}^t\phi \wedge h * \bar{\psi} \tag{7.1}$$

where h is the metric in $\boldsymbol{E}$. Stoke's formula gives

$$(\bar{\partial}\phi, \psi) - (\phi, \vartheta\psi) = \int_{\partial W} {}^t\phi \wedge h * \bar{\psi} \qquad (\phi \in \boldsymbol{A}^{q-1}(\boldsymbol{E}), \psi \in \boldsymbol{A}^q(\boldsymbol{E})). \tag{7.2}$$

We let $\boldsymbol{D}^q(\boldsymbol{E}) \subset \boldsymbol{A}^q(\boldsymbol{E})$ be the domain of the adjoint ϑ of $\bar{\partial}$ (ϑ is given by (6.20)). By (7.2),

$$\boldsymbol{D}^q(\boldsymbol{E}) = \{\psi \in \boldsymbol{A}^q(\boldsymbol{E}) : {}^t\phi \wedge h * \bar{\psi} \mid \partial W = 0 \text{ for all } \phi \in \boldsymbol{A}^{q-1}(\boldsymbol{E})\}. \tag{7.3}$$

For $\phi, \psi \in \boldsymbol{D}^q(\boldsymbol{E})$, we set

$$\boldsymbol{D}(\phi, \psi) = (\bar{\partial}\phi, \bar{\partial}\psi) + (\vartheta\phi, \vartheta\psi) + (\phi, \psi)\,; \tag{7.4}$$

$$\boldsymbol{D}(\phi)^2 = \boldsymbol{D}(\phi, \phi)\,; \tag{7.5}$$

and we let $\boldsymbol{H}^q(\boldsymbol{E}) = \{\phi \in \boldsymbol{D}^q(\boldsymbol{E}) : \bar{\partial}\phi = \vartheta\phi = 0\}$.

THEOREM [15]. *Assume that, for some* q, *we can find* $c > 0$ *such that, for all* $\phi \in \boldsymbol{D}^q(\boldsymbol{E})$, *we have*

$$\boldsymbol{D}(\phi)^2 \geqq c \int_{\partial W} |\phi|^2. \tag{7.6}$$

Then $\dim \boldsymbol{H}^q(\boldsymbol{E}) < \infty$ *and there exists an isomorphism*

$$H^q(W, \Omega(\boldsymbol{E})) \cong \boldsymbol{H}^q(\boldsymbol{E}). \tag{7.7}$$

2. Statement of the vanishing theorems. Let $\boldsymbol{E} \to W$ be as in VII. 1 and let $\boldsymbol{L} \to W$ be a line bundle in which we have a finite metric. Set $\boldsymbol{E}(\mu) = \boldsymbol{E} \otimes \boldsymbol{L}^\mu$ and let $\Theta(\mu) \in \boldsymbol{A}^{1,1}(\operatorname{Hom}(\boldsymbol{E}(\mu), \boldsymbol{E}(\mu)))$ be the curvature form in $\boldsymbol{E}(\mu)$. Our main result is

THEOREM VI. *There exists $c > 0$, which is independent of μ, such that for $\phi \in \boldsymbol{D}^q(\boldsymbol{E}(\mu))$ we have*

$$\boldsymbol{D}(\phi, \phi) \geqq c\{-\int_W \Theta(\mu)^{\#}(\phi, \phi) - \int_{\partial W} L^{\#}(\phi, \phi)\} \qquad (0 \leqq q < n) \tag{7.8}$$

$$\boldsymbol{D}(\phi, \phi) \geqq c\{\int_W \Theta(\mu)(\phi, \phi) + \int_{\partial W} L(\phi, \phi)\} \quad (0 < q \leqq n). \tag{7.9}$$

(*Remark.* The Hermitian forms $\Theta(\mu)(\phi, \phi)$, $\Theta(\mu)^{\#}(\phi, \phi)$, $L(\phi, \phi)$, and $L^{\#}(\phi, \phi)$ are those discussed in § VI. 4.)

From Theorem VI we shall prove

THEOREM 7. 1. *Suppose that there exist integers s, t and positive constants μ_0, λ such that, for all $\phi \in \boldsymbol{D}^q(\boldsymbol{E}(\mu))$ with $\mu \geqq \mu_0$, we have*

$$\begin{cases} \Theta(\mu)^{\#}(\phi, \phi) \leqq -\lambda\mu\,|\,\phi\,|^2 \text{ and} \\ L^{\#}(\phi, \phi) \leqq -\lambda\,|\,\phi\,|^2 \quad \text{for } 0 \leqq q < s; \text{ and} \end{cases} \tag{7.10}$$

$$\begin{cases} \Theta(\mu)(\phi, \phi) \geqq \lambda\mu\,|\,\phi\,|^2 \text{ and} \\ L(\phi, \phi) \geqq \lambda\,|\,\phi\,|^2 \quad \text{for } n - t < q \leqq n. \end{cases} \tag{7.11}$$

Then there exists a v_0 such that

$$H^q(W, \Omega(\boldsymbol{E}(\mu))) = 0 \text{ for } \mu \geqq v_0,\ 0 \leqq q < s,\ n - t < q \leqq n. \tag{7.12}$$

Proof of Theorem 7.1 *from Theorem* IV. From (7.8) and (7.10) we get, for $\phi \in \boldsymbol{D}^q(\boldsymbol{E})$ $0 \leqq q < s$, the estimate

$$\boldsymbol{D}(\phi)^2 \geqq c\lambda\mu \int_W |\,\phi\,|^2 + c\lambda \int_{\partial W} |\,\phi\,|^2. \tag{7.13}$$

From (7.13) and (7.6), we have $H^q(W, \Omega(\boldsymbol{E}(\mu))) \cong \boldsymbol{H}^q(\boldsymbol{E}(\mu))$ and, for $\phi \in \boldsymbol{H}^q(\boldsymbol{E}(\mu))$, we get $(c\lambda\mu - 1) \int_W |\,\phi\,|^2 \leqq 0$. Thus, for $\mu > \frac{1}{c\lambda}$, $\boldsymbol{H}^q(\boldsymbol{E}(\mu)) = 0$ and $H^q(W, \Omega(\boldsymbol{E}(\mu))) = 0$ for $\mu > \frac{1}{c\lambda}$, $0 \leqq q < s$. From (7.9) and

(7.11) we similarly find $H^q(W,\Omega(\boldsymbol{E}(\mu)))=0$ for $\mu\geqq\frac{1}{c\lambda}$, $n-t<q\leqq n$. Q.E.D.

The rest of Section VII will be devoted to proving Theorem VI.

3. A special integral formula. We shall prove:

THEOREM 7.2. *For* $\phi\in\boldsymbol{D}^q(\boldsymbol{E})$, we have

$$(\bar\partial\phi,\bar\partial\phi)+(\vartheta\phi,\vartheta\phi)=(\bar\nabla\phi,\bar\nabla\phi)+\int_W\Theta(\phi,\phi)+\int_W\Phi(\phi,\phi)+\int_{\partial W}L(\phi,\phi)+\int_W\Sigma(\phi,\phi)\tag{7.14}$$

where

$$\Sigma(\phi,\phi)=2Re\langle[\bar\partial,T^{\#}]\phi,\phi\rangle-\langle[T,T^{\#}]\phi,\phi\rangle.\tag{7.15}$$

Proof. This is a routine calculation and we shall only give the relevant formulas. First we have

$$(\bar\partial-T)(\phi)=\frac{(-1)^q}{(q+1)!}\sum_J(\phi_{\bar j_1\cdots\bar j_q|\bar j_{q+1}}-\phi_{\bar j_{q+1}\bar j_2\cdots\bar j_q|\bar j_1}-\cdots-\phi_{\bar j_1\cdots\bar j_{q-1}\bar j_{q+1}|\bar j_q})\overline{\omega}^J\tag{7.16}$$

and

$$(\vartheta-T^{\#})\phi)=\frac{1}{(q-1)!}\sum_{I,j}\phi_{\bar i_1\cdots\bar i_{q-1}\bar j}\,{}_{|\bar j_q}\overline{\omega}^I.\tag{7.17}$$

Set $\xi=\frac{1}{(q-1)!}\sum\phi_{\bar j_1\cdots\bar j_{q-1}\bar j^{\rho}|j}\bar\phi_{\bar j_1\cdots\bar j_{q-1}\bar k^{\rho}}\omega^k$. Then ξ is a globally defined (1,0) form and

$$-\delta\xi=\langle(\vartheta-T^{\#})\phi,(\vartheta-T^{\#})\phi\rangle+\frac{1}{(q-1)!}\sum\phi_{\bar j_1\cdots\bar j_{q-1}\bar j^{\rho}|j|\bar k}\bar\phi_{\bar j_1\cdots\bar j_{q-1}\bar k^{\rho}}.\tag{7.18}$$

By (7.18) and Green's theorem we have

$$((\vartheta-T^{\#})\phi,(\vartheta-T^{\#})\phi)=-\frac{1}{(q-1)!}\int_W\phi_{\bar j_1\cdots\bar j_{q-1}\bar j^{\rho}|j|\bar k}\bar\phi_{\bar j_1\cdots\bar j_{q-1}\bar k}+\int_{\partial W}*\bar\xi.\tag{7.19}$$

If, in a neighborhood of ∂W, we choose $\omega^1,\cdots,\omega^n$ as in VI.1, then the boundary conditions (7.3) become

$$\phi_I{}^{\rho} \,|\, \partial W = 0 \text{ whenever } n \in I. \tag{7.20}$$

Thus, on $\partial W, *\xi = \frac{1}{(q-1)!} \sum (\phi_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{j}}{}^{\rho} \,|\, {}_j \phi_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{n}}{}^{\rho}) *\omega^n = 0$. The Ricci identity gives

$$\begin{aligned} \frac{1}{(q-1)!}\{\sum {}_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{j} \,|\, \bar{k} \,|\, j}{}^{\rho} \bar{\phi}_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{k}}{}^{\rho} - \phi_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{j} \,|\, j \,|\, \bar{k}}{}^{\rho} \bar{\phi}_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{k}}{}^{\rho} \\ = \Theta(\phi, \phi) + \Phi(\phi, \phi). \end{aligned} \tag{7.21}$$

Combining (7.19) and (7.21) gives

$$\begin{aligned} ((\vartheta - T^{\#})\phi, (\vartheta - T^{\#})\phi) &= \int_W \Theta(\phi, \phi) + \int_W \Phi(\phi, \phi) \\ &\quad - \frac{1}{(q-1)!} \int_W \{\sum \phi_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{j}}{}^{\rho} {}_{|\, \bar{k} \,|\, j} \bar{\phi}_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{k}}{}^{\rho}\}. \end{aligned} \tag{7.22}$$

Now let $\eta = \frac{1}{(q-1)!} \phi_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{j} \,|\, \bar{k}}{}^{\rho} \bar{\phi}_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{k}}{}^{\rho} \overline{\omega}^j$. Then η is a (0,1) form on W and

$$\begin{aligned} \delta\eta = -\frac{1}{(q-1)!}\{\sum \phi_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{j} \,|\, \bar{k} \,|\, j}{}^{\rho} \bar{\phi}_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{k}}{}^{\rho} \\ + \phi_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{j} \,|\, \bar{k}}{}^{\rho} \bar{\phi}_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{k} \,|\, \bar{j}}{}^{\rho}\}. \end{aligned} \tag{7.23}$$

From the algebraic identity (using (7.16))

$$\begin{aligned} -\frac{1}{(q-1)!} \sum \phi_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{j} \,|\, \bar{k}}{}^{\rho} \bar{\phi}_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{k} \,|\, \bar{j}}{}^{\rho} = -\langle(\bar{\partial} - T)\phi, (\bar{\partial} - T)\phi\rangle \\ + \langle\bar{\nabla}\phi, \bar{\nabla}\phi\rangle \end{aligned}$$

and (7.22), (7.23), we get

$$\begin{aligned} ((\vartheta - T^{\#})\phi, (\vartheta - T^{\#})\phi + ((\bar{\partial} - T)\phi, (\bar{\partial} - T)\phi) \\ = (\bar{\nabla}\phi, \bar{\nabla}\phi) + \int_W \Theta(\phi, \phi) + \int_W \Phi(\phi, \phi) + \int_{\partial W} *\eta. \end{aligned} \tag{7.24}$$

Thus, in order to prove (7.14), we must show

$$*\eta \,|\, \partial W = L(\phi, \phi)\, d\sigma \qquad (d\sigma = \text{area element on } \partial W). \tag{7.25}$$

Now

$$*\eta \,|\, \partial W = \frac{1}{(q-1)!} \sum_{k<n} \phi_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{n} \,|\, \bar{k}}{}^{\rho} \bar{\phi}_{\bar{j}_1 \cdots \bar{j}_{q-1}\bar{k}}{}^{\rho} *\omega^n.$$

By (6.6),

$$\phi_{\bar{j}_1\cdots\bar{j}_{q-1}\bar{n}\,|\,\bar{k}}{}^{\rho} = e_{\bar{k}}\cdot\phi_{\bar{j}_1\cdots\bar{j}_{q-1}\bar{n}}{}^{\rho} - \sum \bar{\Lambda}_{nk}{}^{j}\phi_{\bar{j}_1\cdots\bar{j}_{q-1}\bar{j}}{}^{\rho}$$

$$-\sum_{s<q}\bar{\Lambda}_{j_s k}{}^{j}\phi_{\bar{j}_1\cdots(\bar{j})_s\cdots\bar{j}_{q-1}\bar{n}}.$$

Since $e_1,\cdots,e_{n-1}$ are tangent to ∂W, and $\phi_{\bar{j}_1\cdots\bar{j}_{q-1}\bar{n}}{}^{\rho}\,|\,\partial W = 0$, we have an ∂W the identity

$$\sum\phi_{\bar{j}_1\cdots\bar{j}_{q-1}\bar{n}\,|\,\bar{k}}{}^{\rho}\bar{\phi}_{\bar{j}_1\cdots\bar{j}_{q-1}\bar{k}}{}^{\rho} = -\sum_{j,k<n}\Lambda_{nk}{}^{j}\phi_{\bar{j}_1\cdots\bar{j}_{q-1}\bar{j}}{}^{\rho}\bar{\phi}_{\bar{j}_1\cdots\bar{j}_{q-1}\bar{k}}{}^{\rho}.$$

From (6.5) and (6.14) it follows that $*\eta\,|\,\partial W = L(\phi,\phi)\,d\sigma$. Q. E. D.

Proof of (7.9) *in Theorem VI.* First of all, we may find $\alpha > 0$ such that, for all $\epsilon > 0$, we have

$$|\,2Re\int_W\langle[T^{\#},\bar{\partial}]\phi,\phi\rangle\,| \leqq \alpha\epsilon(\bar{\nabla}\phi,\bar{\nabla}\phi) + \frac{\alpha}{\epsilon}(\phi,\phi)$$

and

$$|\int_W\Phi(\phi,\phi)\,| + |\int_W\langle[T,T^{\#}]\phi,\phi\rangle\,| \leqq \alpha(\phi,\phi).$$

Combining these inequalities with (7.14), (7.15) gives

$$(\bar{\partial}\phi,\bar{\partial}\phi) + (\vartheta\phi,\vartheta\phi) \geqq (1-\alpha\epsilon)(\bar{\nabla}\phi,\bar{\nabla}\phi) + \int_W\Theta(\phi,\phi)$$

$$+\int_{\partial W}L(\phi,\phi) - (\frac{\alpha}{\epsilon}+\alpha)(\phi,\phi).$$

If we choose $\epsilon < 1/\alpha$ and $c < \dfrac{\epsilon}{\alpha+\alpha\epsilon}$, (7.9) is satisfied. Indeed, it is clear from our computations that the constant c in (7.9) may be chosen to be independent of μ.

4. Some remarks and applications.

(i) Compact manifolds. Let X be a compact Hermitian manifold and $\boldsymbol{E}\to X$ a holomorphic vector bundle in which we have an Hermitian metric.

PROPOSITION 7.1. *If, for some q, the Hermitian form*

$$\Theta(\phi,\phi) + \Phi(\phi,\phi) - |\,T\phi\,|^2 - |\,T^{\#}\phi\,|^2 \tag{7.26}$$

is positive definite on all of X ($\phi\in A^q(\boldsymbol{E})$), then $H^q(X,\Omega(\boldsymbol{E})) = 0$.

Proof. $H^q(X, \Omega(\boldsymbol{E})) \cong \boldsymbol{H}^q(\boldsymbol{E})$ and, for $\phi \in \boldsymbol{H}^q(\boldsymbol{E})$,

$$\int_X \{\Theta(\phi,\phi) + \Phi(\phi,\phi) - |T\phi|^2 - |T^{\#}\phi|^2\} \leqq 0$$

by Theorem 6.2. Thus, if $\phi \in \boldsymbol{H}^q(\boldsymbol{E})$ and if (7.26) is positive definite, we must have $\phi \equiv 0$.

Now $\Theta(\phi,\phi) + \Phi(\phi,\phi) = \Theta_{\boldsymbol{E}\otimes\boldsymbol{K}^*}(\phi,\phi)$ where $\boldsymbol{K}$ is the canonical bundle. Thus (7.26) is

$$\Theta_{\boldsymbol{E}\otimes\boldsymbol{K}^*}(\phi,\phi) - \{|T\phi|^2 + |T^{\#}\phi|^2\} \tag{7.27}$$

((7.27) shows clearly the effect on the curvature of changing the metric and perhaps adding torsion.) From (7.27) and Proposition 7.1, we get the following:

a. If X is Kähler and $\boldsymbol{E}\otimes\boldsymbol{K}^*$ is positive, then $H^q(X, \Omega(\boldsymbol{E})) = 0$ for q > 0 (Kodaira vanishing theorem).

b. If X is Kähler and if $\boldsymbol{E}$ has signature $(0,t)$, then $H^q(X,\Omega(\boldsymbol{E})) = 0$ for $q < t$ (for $t = n$, this reduces to *a*).

c. If $\boldsymbol{E} \to X$ has signature (s,t), then, for $\mu \geqq \mu_0$,

$$H^q(X, \Omega(\boldsymbol{E}^\mu)) = 0 \text{ for } 0 \leqq q < t,\ n - s < q \leqq n. \tag{7.28}$$

(This is the vanishing theorem of [3], §23.)

Proof. In case $\boldsymbol{E}$ is a line bundle, we use Proposition 6.4 and choose a metric on X such that $\Theta(\phi,\phi)$ is positive definite on $A^q(\boldsymbol{E})$ for $n - s < q \leqq n$. Then, for μ sufficiently large,

$$\Theta_{\boldsymbol{E}^\mu\otimes\boldsymbol{K}^*}(\phi,\phi) - \{|T\phi|^2 + |T^{\#}\phi|^2\}$$

will be positive definite for $n - s < q \leqq n$ and we may use Proposition 7.1 plus Serre duality to prove *c.* for line bundles. The general case follows from the case of line bundles, Proposition 6.5, and the fact that

$$H^q(X, \Omega(\boldsymbol{E}^\mu) \cong H^q(P(\boldsymbol{E}^*), \Omega(\boldsymbol{L}^\mu)).$$

(ii) Open manifolds. Let $W \subset W'$ be a finite complex manifold. We say that W is *t-complete* if there exists a real function r on W' such that $\{r < 0\}$ defines $W \subset W'$ and $\partial\bar{\partial}r$ has everywhere t positive eigenvalues.

PROPOSITION 7.2. *If W is t-complete and $\boldsymbol{E} \to W$ is any holomorphic vector bundle, then*

$$(7.29)\qquad H^q(W, \Omega(\boldsymbol{E})) = 0 \text{ for } q > n - t.$$

Proof. If $\boldsymbol{L} \to W$ is the trivial line bundle, then e^{-r} is a metric in $\boldsymbol{L}$ whose curvature is $\bar{\partial}\partial \log e^{-r} = \partial\bar{\partial} r$. Thus (7.11) will hold for large μ, and

so $0 = H^q(W, \Omega(\boldsymbol{E}(\mu))) = H^q(W, \Omega(\boldsymbol{E}))$ for $n - t < q \leqq n$.

Remark. For $t = n$ this is closely related to Cartan's Theorem B (for locally free sheaves). A more general form of (7.29) has been given by Andreotti-Vesentini [2].

5. An integral formula for functions on complex manifolds. The integral formula (7.14) was based entirely on a formula in Riemannian geometry (Bochner-Yano [4]) and, in fact, the precise analogue of (7.14) (boundary values included and with the second fundamental form of $\partial W \subset W'$ replacing the Levi form) has been in existence for some years (cf. [11]). The next formula we shall give has no Riemannian analogue and is based on the existence of an $(n-1, n-1)$ form Ω on W which has the property

$$(7.30)\qquad (\partial - \bar{\partial})\Omega \mid \partial W = 2(\operatorname{Trace} L)\sigma.$$

For example, in $\boldsymbol{C}^n$ with coordinates $z = (z^1, \cdots, z^n)$,

$$\Omega = (-1)^{\frac{n(n+1)}{2}} \left(\frac{\sqrt{-1}}{2}\right)^n \sum \left\{ \frac{(-1)^{i+j} z^i \bar{z}^j}{\sum |z^k|^2} - \delta_j{}^i \right\}$$
$$(dz^1 \cdots \widehat{dz^i} \cdots dz^n d\bar{z}^1 \cdots \widehat{d\bar{z}^j} \cdots d\bar{z}^n).$$

(we shall omit the exterior product symbols in this paragraph). For simplicity, we shall assume that the Hermitian metric on W' is without torsion. In the general case, the terms with torsion can be absorbed in estimates as in the proof of (7.9) given in § VII. 3 above. Let $\gamma = \frac{\sqrt{-1}}{2} \sum \omega^i \wedge \bar{\omega}^i$ be the fundamental form of the Kähler metric and consider real non-negative functions τ on W' which are equal to one in a neighborhood of ∂W. For each such τ, we define the $(n-1, n-1)$ form Ω_τ by

$$(7.31)\qquad \Omega_\tau = \gamma^{n-1} - \tau * (\partial r \wedge \overline{\partial r})$$

Theorem 7.3. *For g a C^∞ function on $\bar{W}$, we have*

$$(7.32)\qquad \int_W g\, \bar{\partial}\partial \Omega_\tau + \int_W (\partial\bar{\partial} g) \wedge \Omega_\tau = \int_{\partial W} g (\operatorname{Trace} L)\sigma.$$

We begin with some preliminary computations. If

$$\lambda = (-1)^{\frac{n(n-1)}{2}} \left(\frac{\sqrt{-1}}{2}\right)^n,$$

then the volume on W is given by $\Psi = \lambda \omega^1 \wedge \cdots \wedge \omega^n \wedge \bar{\omega}^1 \wedge \cdots \wedge \bar{\omega}^n$. We let

$$\Psi_k = (-1)^{k-1} \lambda (\omega^1 \cdots \hat{\omega}^k \cdots \omega^n \cdots \bar{\omega}^1 \cdots \bar{\omega}^n)$$

and

$$\Psi_{\bar{k}} = (-1)^{n+k-1} \lambda (\omega^1 \cdots \omega^n \cdots \bar{\omega}^1 \cdots \hat{\bar{\omega}}^k \cdots \bar{\omega}^n).$$

Then $\Psi_k = {*\bar{\omega}^k}$ and $\Psi_{\bar{k}} = {*\omega^k}$. Also we set

$$\Psi_{i\bar{j}} = (-1)^{n+i+j} \lambda (\omega^1 \cdots \hat{\omega}^i \cdots \omega^n \bar{\omega}^1 \cdots \hat{\bar{\omega}}^j \cdots \bar{\omega}^n).$$

Since $| dr |^2 = | \omega^n |^2 + | \bar{\omega}^n |^2 = 2$, $\sigma = (\tfrac{1}{2} {*dr}) \mid \partial W$. Thus

$$\sigma = \tfrac{1}{2}(\Psi_{\bar{n}} + \Psi_n) \mid \partial W = \Psi_{\bar{n}} \mid \partial W$$

since $(\omega^n + \bar{\omega}^n) \mid \partial W = 0$. Using the fact that the metric has no torsion, a computation based on (6.2), (6.4), (6.5) gives

$$d\Psi_j = (\sum \bar{\Lambda}_{ii}{}^j) \Psi \tag{7.33}$$

$$d\Psi_{\bar{j}} = (\sum \Lambda_{i\bar{i}}{}^j) \Psi. \tag{7.34}$$

Since $\gamma^{n-1} = (-1)^{n-1} (\sum_{k=1}^{n} \Psi_{k\bar{k}})$ we find that, in a neighborhood of ∂W,

$$\Omega_T = (-1)^{n-1} (\sum_{k<n} \Psi_{k\bar{k}}). \tag{7.35}$$

LEMMA 7.1. $(\partial - \bar{\partial}) \Omega_T \mid \partial W = +2 (\operatorname{Trace} L) \sigma$.

Proof. Since $\omega^n \wedge \bar{\omega}^n \mid \partial W = 0$, we shall neglect terms containing $\omega^n \wedge \bar{\omega}^n$. Then

$$\begin{aligned}(-1)^{n-1} \partial \Psi_{k\bar{k}} &= (-1)^n (-1)^{2n-3} \lambda (\omega^1 \cdots \hat{\omega}^k \cdots \omega^n \bar{\omega}^1 \cdots \hat{\bar{\omega}}^k \cdots \partial \bar{\omega}^n) \\ &= (-1)^{n+1} \lambda \bar{\Lambda}_{k\bar{k}}{}^n (\omega^1 \cdots \hat{\omega}^k \cdots \omega^n \bar{\omega}^1 \cdots \hat{\bar{\omega}}^k \cdots \bar{\omega}^{n-1} \omega^k \bar{\omega}^k) \\ &= \lambda \Lambda_{k\bar{k}}{}^n (\omega^1 \cdots \omega^n \bar{\omega}^1 \cdots \bar{\omega}^{n-1}) = \Lambda_{k\bar{k}}{}^n \Psi_{\bar{n}}.\end{aligned}$$

Similarly, $-(1)^{n-1} \bar{\partial} \Psi_{k\bar{k}}{}^n = \Lambda_{k\bar{k}}{}^n \Psi_{\bar{n}}$. Thus

$$(\partial - \bar{\partial}) \Omega \mid \partial W = 2 (\operatorname{Trace} L) \Psi_{\bar{n}} \mid \partial W = 2 (\operatorname{Trace} L) \sigma.$$

We can now complete the proof of (7.32). Lemma 7.1 and Stokes theorem give

$$+\int_{\partial W} g(\mathrm{Trace}\,L)\sigma = \tfrac{1}{2}\int_W d\{g(\partial-\bar\partial)\Omega_\tau\}. \tag{7.36}$$

Now

$$\begin{aligned} d\{g(\partial-\bar\partial)\Omega_\tau\} &= 2g\bar\partial\partial\Omega_\tau + (dg)\wedge(\partial-\bar\partial)\Omega_\tau \\ &= 2g\bar\partial\partial\Omega_\tau + d\{(\partial g)\wedge\Omega_\tau\} - d\{(\bar\partial g)\wedge\Omega_\tau + 2\partial\bar\partial g\wedge\Omega_\tau\}. \end{aligned}$$

By (7.36) and Stokes theorem again (since $\Omega_\tau \mid \partial W = 0$), we get

$$+\int_{\partial W} g(\mathrm{Trace}\,L) = \int_W g\bar\partial\partial\Omega_\tau + \int_W \partial\bar\partial g\wedge\Omega_\tau. \quad \text{Q.E.D.}$$

We now define operators $\bar\partial_\tau$, ∂_τ, and Δ_τ by

$$\bar\partial_\tau f = \bar\partial f - \tau\langle\bar\partial f, \bar\partial r\rangle\bar\partial r, \qquad \partial_\tau f = \partial f - \tau\langle\partial f, \partial r\rangle\partial r,$$

$$\Delta_\tau f = \langle\partial\bar\partial f, \gamma - \tau\partial r\wedge\bar\partial r\rangle.$$

At those points where $\tau = 0$, $\bar\partial_\tau f = \bar\partial f$, $\partial_\tau f = \partial f$, and $\Delta_\tau f = \Delta f$. From (7.32) we get, upon setting $g = |f|^2$ and $\Phi_\tau = \bar\partial\partial\Omega_\tau$,

$$\begin{aligned} &\int_W |f|^2\Phi_\tau + \int_W |\partial_\tau f|^2 + 2Re\int_W \Delta_\tau f\cdot\bar f \\ &\qquad = \int_W |\bar\partial_\tau f|^2 + \int_{\partial W} |f|^2(\mathrm{Trace}\,L). \end{aligned} \tag{7.37}$$

In particular, if f is holomorphic, then

$$\int_W |f|^2\Phi_\tau + \int_W |\partial_\tau f|^2 = \int_{\partial W} |f|^2(\mathrm{Trace}\,L) \tag{7.38}$$

which is, for us, an important formula. For example, suppose that L has one negative eigenvalue and choose a metric so that $\mathrm{Trace}\,L < 0$. Then we have $\int_W |f|^2\Phi_\tau + \int_W |\partial_\tau f|^2 \leqq 0$. One may prove that this inequality implies that f is constant.

Example. We shall give an integral formula in $\boldsymbol{C}^n$ which is similar to, but does not quite follow from, (7.38). Let B = unit ball in $\boldsymbol{C}^n$, $S = \partial B$. Let

$$\omega = \lambda \Big(\sum_{k=1}^{n} (-1)^{n+1} (\cdots \widehat{dz^k} \cdots \widehat{d\bar{z}^k} \cdots) \Big)$$

(here $\quad (\cdots \widehat{dz^k} \cdots \widehat{d\bar{z}^k} \cdots)$

$$= (dz^1 \wedge \cdots \wedge \widehat{dz^k} \wedge \cdots \wedge d\bar{z}^1 \wedge \cdots \wedge \widehat{d\bar{z}^k} \wedge \cdots \wedge d\bar{z}^n))$$

and set

$$\Omega = \omega + \lambda \sum (-1)^{i+n+j} z^i \bar{z}^j (\cdots \widehat{dz^i} \cdots \widehat{d\bar{z}^j} \cdots).$$

Then $(\partial - \bar{\partial})\Omega = \lambda \{ \sum (-1)^{i-1} z^i (\cdots \widehat{dz^i} \cdots) + \sum (-1)^{n+j-1} \bar{z}^j (\cdots \widehat{d\bar{z}^j} \cdots) \}$ so that $d(\partial - \bar{\partial})\Omega = 2n\Psi$ where Ψ is the volume element. We claim that $\Omega \mid S = 0$. Since Ω is invariant under the unitary group acting on S, it will suffice to show that $(\Omega \mid S)(x_0) = 0$ where $x_0 = (1, 0, \cdots, 0)$. On S, $\sum \bar{z}^i dz^i + \sum z^j d\bar{z}^j = 0$ so that, at x_0 $dz^1 + d\bar{z}^1 = 0$. But then $(\omega \mid S)(x_0) = \lambda(-1)^{n+1} (\widehat{dz^1} \cdots \widehat{d\bar{z}^1} \cdots)$ and so $(\Omega \mid S)(x_0) = 0$. Similarly we find that $(\partial - \bar{\partial})\Omega \mid S = 2\sigma$ where σ is the area element on S. Using Stoke's theorem twice as in the proof of Theorem 7.3, we find the formula

$$\int_B \{ \sum | \frac{\partial f}{\partial z^k} |^2 - | \sum z^k \frac{\partial f}{\partial z^k} |^2 \} + n \int_B | f |^2 = \int_S | f |^2. \tag{7.39}$$

For example, on the unit disc writing $f(z) = \sum a_n z^n$, (7.39) reduces to the identity

$$2\pi \{ \sum \frac{n^2 | an |^2}{2n} - \sum \frac{n^2 | an |^2}{2n+2} + \sum \frac{| an |^2}{2n+2} \} = \pi \{ \sum | an |^2 \}.$$

6. Proof of Theorem VI. Let $\boldsymbol{F} \to W$ be a holomorphic vector bundle and let $\phi \in \boldsymbol{D}^q(\boldsymbol{F})$. In the same way as we defined $\bar{\partial}_\tau f$ and $\partial_\tau f$ above, we may now define $\bar{\bar{\nabla}}^\tau \phi$ and $\bar{\nabla}^\tau \phi$. Writing $\bar{\nabla} \phi = \bar{\bar{\nabla}}^\tau \phi + \bar{\nabla}^\eta \phi$ ($\bar{\nabla}^\eta =$ covariant differential in the normal direction), we have

$$| \bar{\nabla} \phi |^2 = | \bar{\bar{\nabla}}^\tau \phi |^2 + | \bar{\bar{\nabla}}^\eta \phi |^2. \tag{7.40}$$

Also, we define $\operatorname{Trace} \Theta^\tau (\phi, \phi) = * \{ \frac{1}{q!} \sum \Theta_{\sigma i j}{}^{\rho} \phi_{\bar{I}}{}^{\sigma} \bar{\phi}_{\bar{I}}{}^{\rho} \, \omega^i \wedge \bar{\omega}^j \wedge \Omega_\tau \}$. At those points where $\tau = 0$, $\operatorname{Trace} \Theta^\tau (\phi, \phi) = \operatorname{Trace} \Theta (\phi, \phi)$ as given in (6.37). Our proof of (7.8) is based on the following:

THEOREM 7.4. *If the metric is without torsion, then*

$$(\bar{\partial}\phi, \bar{\partial}\phi) + (\vartheta\phi, \vartheta\phi) \geqq (\bar{\bar{\nabla}}^\eta \phi, \bar{\nabla}^\eta \phi) + (\nabla^\tau \phi, \nabla^\tau \phi)$$

$$-\int_W \Theta^\#(\phi,\phi) - \int_{\partial W} L^\#(\phi,\phi) + \int_W \Xi(\phi,\phi) \tag{7.41}$$

where

$$\begin{aligned}\Xi(\phi,\phi) &= \Phi(\phi,\phi) + |\phi|^2\,|\Phi_\tau|^2 \\ &\quad + \{\text{Trace}\,\Theta(\phi,\phi) - \text{Trace}\,\Theta^\tau(\phi,\phi)\}\end{aligned} \tag{7.42}$$

Proof. We shall prove (7.41) by combining (7.14) and (7.32) when $g = |\phi|^2$. Now

$$\begin{aligned}\partial\bar\partial\,|\phi|^2 = \frac{1}{q!}\{&2Re\sum \phi_{\bar I^\rho\,|\,\bar k\,|\,j}\bar\phi_{\bar I^\rho} \\ &+ \sum \phi_{\bar I\,|\,j^\rho}\bar\phi_{\bar I\,|\,k^\rho} - \sum \phi_{\bar I\,|\,\bar k^\rho}\bar\phi_{\bar I\,|\,\bar j k} \\ &- \sum \Theta_{\sigma j k^\rho}\phi_{\bar I^\sigma}\bar\phi_{\bar I^\rho}\}\omega^j \wedge \omega^k.\end{aligned}$$

Thus

$$\begin{aligned}\partial\bar\partial\,|\phi|^2 \wedge \Omega_\tau = \{&|\nabla^\tau\phi|^2 - |\bar\nabla^\tau\phi|^2 \\ &- \text{Trace}\,\Theta^\tau(\phi,\phi) + 2Re\langle\nabla_{k^\tau}\bar\nabla_{k^\tau}\phi,\phi\rangle\}\Psi.\end{aligned} \tag{7.43}$$

Because the operator ∇^τ involves no normal derivatives, we may integrate by parts to find

$$\int_W \{2Re\sum\langle\nabla_{k^\tau}\bar\nabla_{k^\tau}\phi,\phi\rangle\} \leqq 0. \tag{7.44}$$

From (7.43) and (7.44) we get

$$\begin{aligned}\int_W \partial\bar\partial\,|\phi|^2 \wedge \Omega_\tau \geqq \int_W |\nabla^\tau\phi|^2 - \int_W |\bar\nabla^\tau\phi|^2 \\ - \int_W \text{Trace}\,\Theta^\tau(\phi,\phi).\end{aligned} \tag{7.45}$$

We now use (7.32) in the form

$$\int_W |\phi|^2\Phi_\tau - \int_{\partial W}(\text{Trace}\,L)\,|\phi|^2 = -\int_W \partial\bar\partial\,|\phi|^2 \wedge \Omega_\tau. \tag{7.46}$$

Combining (7.45) and (7.46) gives

$$\begin{aligned}\int_W |\bar\nabla^\tau\phi|^2 \geqq \int_W |\nabla^\tau\phi|^2 - \int_W \text{Trace}\,\Theta^\tau(\phi,\phi) \\ - \int_{\partial W}(\text{Trace}\,L)\,|\phi|^2 + \int_W |\phi|^2\Phi_\tau.\end{aligned} \tag{7.47}$$

$$\begin{aligned}(\bar{\partial}\phi, \bar{\partial}\phi) + (\vartheta\phi, \vartheta\phi) = (\bar{\nabla}^{\tau}\phi, \bar{\nabla}^{\tau}\phi) + (\bar{\nabla}^{\eta}\phi, \bar{\nabla}^{\eta}\phi) \\ + \int_{W} \Theta(\phi, \phi) + \int_{\partial W} L(\phi, \phi).\end{aligned} \tag{7.48}$$

Then (7.48) and (7.47) taken together give (7.41). Q.E.D.

Now in case the metric has torsion, we must, at least, add the term $2Re \int_{W} \langle [\bar{\partial}, T^{\#}]\phi, \phi\rangle - \int_{W} \langle [T, T^{\#}]\phi, \phi\rangle$ to the right hand side of (7.41). Since $\phi \in \boldsymbol{D}^{q}(\boldsymbol{F})$

$$\int_{W} \langle [\bar{\partial}, T^{\#}]\phi, \phi\rangle = \int_{W} \langle [\bar{\partial}_{\eta}, T^{\#}]\phi, \phi\rangle + \int_{W} \langle [\vartheta_{\tau}, T]\phi, \phi\rangle$$

and we then get an estimate

$$\left| 2Re \int_{W} \langle [\bar{\partial}, T^{\#}]\phi, \phi\rangle \right| \leqq \epsilon\beta(\bar{\nabla}^{\eta}\phi, \bar{\nabla}^{\eta}\phi) + \epsilon\beta(\nabla^{\tau}\phi, \nabla^{\tau}\phi) + \frac{\beta}{\epsilon}(\phi, \phi).$$

Thus any such terms involving the torsion can be absorbed in the right hand side of (7.41) and we find an estimate

$$\begin{aligned}(\bar{\partial}\phi, \bar{\partial}\phi) + (\vartheta\phi, \vartheta\phi) \geqq - \int_{W} \Theta^{\#}(\phi, \phi) - \int_{\partial W} L^{\#}(\phi, \phi) \\ - \alpha \int_{W} |\phi|^{2} + \int_{W} \{\text{Trace}\,\Theta(\phi, \phi) - \text{Trace}\,\Theta^{\tau}(\phi, \phi)\}\end{aligned} \tag{7.49}$$

where $\alpha > 0$ is a large constant not depending on the bundle $\boldsymbol{F}$. Consequently, in order to prove (7.8) and with it Theorem VI, it will suffice to show that, with an appropriate choice of metrics in $\boldsymbol{E}$ and $\boldsymbol{L}$, we may have

$$\text{Trace}\,\Theta(\mu)(\phi, \phi) - \text{Trace}\,\Theta(\mu)^{\tau}(\phi, \phi) \geqq 0 \tag{7.50}$$

for all ϕ. Now $\text{Trace}\,\Theta(\mu)(\phi, \phi) - \text{Trace}\,\Theta^{\tau}(\phi, \phi) =$

$$\frac{\tau}{q!} \sum \Theta_{\sigma n n}{}^{\rho}(\mu)\, \phi_{\bar{I}}{}^{\sigma} \bar{\phi}_{\bar{I}}{}^{\rho}. \tag{7.51}$$

On the other hand, for any function ψ on W, we may modify the metric in $\boldsymbol{L}$ by a factor $e^{-\psi}$ and this adds to the curvature a term which is essentially $\partial\bar{\partial}\psi$. From (7.51), in order to have (7.50) satisfied we must choose ψ with $\psi_{|\bar{n}|n} = \frac{1}{\tau}\langle \partial\bar{\partial}\psi, \partial r \wedge \bar{\partial} r\rangle > 0$. It is easy to check that any function ψ which is equal to r in a small open neighborhood of $\partial W \subset W'$ will do.

(We are doing nothing more here than using the principle that for any open complex manifold $W \subset W'$ of dimension $n > 1$, we may choose a function r such that $r < 0$ defines $W \subset W'$ and such that $\partial\bar{\partial}\psi$ has a positive normal part. This is closely related to the fact that, for any bundle $\boldsymbol{F} \rightarrow W$, $H^{n}(W, \Omega(\boldsymbol{F})) = 0$.)

VIII. Properties of Tubular Neighborhoods and the Proof of Theorem III. The main result of this section is:

THEOREM VII. *Let $X \subset W$ be a germ of embedding whose normal bundle $\boldsymbol{N}$ has signature (s,t) $(s+t=n=\dim X)$. Suppose furthermore that X is a hyperpsurface and set $\boldsymbol{L}=[X]$ so that $\Omega(\boldsymbol{L}^*)=\mathfrak{J}$, the ideal sheaf of X. Then there exists an embedding $W \subset W'$, Hermitian metrics on W' and in $\boldsymbol{L} \rightarrow W'$, and a function r on W' with $W=\{r<0\}$ such that:*

(i) *$\Theta^{\#}(\phi,\phi)$ is negative definite for $\phi \in \boldsymbol{A}^{q}(\boldsymbol{L})$ and $L^{\#}(\phi,\phi)$ is negative definite for $\phi \in \boldsymbol{D}^{q}(\boldsymbol{L}^*)$, $0 \leqq q < s$;*

(ii) *$\Theta(\phi,\phi)$ is positive definite for $\phi \in \boldsymbol{A}^{q}(\boldsymbol{L}^*)$ and $L(\phi,\phi)$ is positive definite on $\boldsymbol{D}^{q}(\boldsymbol{L}^*)$, $n-t<q \leqq n+1$.*

Proof of Theorem III. By Theorem VII we see that (7.10) and (7.11) are satisfied and so, by Theorem 7.1, we have

$$H^{q}(W, \Omega(\boldsymbol{E} \otimes \boldsymbol{L}^{*\mu}))=0 \text{ for } \mu \geqq \mu_{0}(\boldsymbol{E}), \quad 0 \leqq q<s, n-t<q \leqq n+1. \tag{8.1}$$

Because of Proposition 5.1, (8.1) gives Theorem III.

1. Metrics in line bundles. We retain the notation and assumptions of Theorem VII.

PROPOSITION 8.1. *For $W \supset X$ sufficiently small, there exists in $\boldsymbol{L}^* \rightarrow W$ an Hermitian metric whose curvature has signature $(t, s+1)$.*

Proof. Relative to a suitable coordinate covering $\{U_i\}$ of W, $X \cap U_i$ is given by a local equation $f_i=0$. The transition functions of $\boldsymbol{L}^*$ are $g_{ij}=f_j/f_i$ and a metric in $\boldsymbol{L}^*$ is given by a family of positive functions ψ_i in U_i which satisfy $\psi_j=|g_{ij}|^2\psi_i$ in $U_i \cap U_j$. We may choose ψ such that $\psi \mid X$ is the metric ϕ in $\boldsymbol{N}^*=\boldsymbol{L}^* \mid X$ whose curvature has signature (t,s). Thus $\psi_i \mid U_i=\phi_i+\rho\tau_i$ where ρ is the distance from X.

We fix $x_0 \in X \cap U_i$ and change holomorphic coordinates in U_i so that

$(d\psi_i)(x_0) = 0$. Since $(d\psi_i)(x_0) = (d\phi_i + \tau_i d\rho)(x_0)$, and since $d\phi_i$ is tangent to X whereas $d\rho$ is normal, it follows that $(d\phi_i)(x_0) = 0 = \tau_i(x_0)$. At x_0, $\bar\partial\partial\psi_i = \bar\partial\partial\phi_i + (\bar\partial\rho \wedge \partial\tau_i - \partial\rho \wedge \bar\partial\tau_i)$. Since $X \subset W$ is a complex submanifold, $\bar\partial\rho$ and $\partial\rho$ are both zero on $\boldsymbol{T}_{x_0}(X)$. Thus $\bar\partial\partial\psi_i \mid \boldsymbol{T}_{x_0}(X) = \bar\partial\partial\phi_i \mid \boldsymbol{T}_{x_0}(X)$ and this implies that, for W sufficiently small, the curvature form for ψ has signature at least (t, s).

Now we observe that, for any $m \geqq 0$, a new metric in $\boldsymbol{L}^*$ is given by functions $\psi_i(m) = \psi_i + m \mid f_i \mid^2$. Furthermore, $\psi_i(m) \mid X \cap U_i = \phi_i$ and, at x_0 chosen as above, $\bar\partial\partial\psi_i(m) = \bar\partial\partial\psi_i - m df_i \wedge \overline{df_i}$. We claim that, for m sufficiently large, $\bar\partial\partial\psi_i$ will have signature $(t, s+1)$.

To prove this, we choose holomorphic coordinates $z_1, \cdots, z_n$; w for W around x_0 such that: (i) $f_i = w$ and $z_1, \cdots, z_n$ are coordinates along X; (ii) $d\psi_i(x_0) = 0$; and (iii)

$$\bar\partial\partial\phi_i(x_0) = -\sum_{1 \leqq \alpha \leqq s} c_\alpha dz^\alpha \wedge d\bar z^\alpha + \sum_{n-t < \rho \leqq n} c_\rho dz^\rho \wedge d\bar z^\rho$$

where $c_\alpha > 0$, $c_\rho > 0$. Let $\boldsymbol{S} \subset \boldsymbol{T}_{x_0}(W)$ be spanned by $\dfrac{\partial}{\partial z^1}, \cdots, \dfrac{\partial}{\partial z^s}; \dfrac{\partial}{\partial w}$. Then $\zeta \in \boldsymbol{S}$ is an $(s+1)$-triple $\zeta = (\lambda_1, \cdots, \lambda_s; \xi)$ and

$$\langle \bar\partial\partial\psi_i(m), \zeta \wedge \bar\zeta \rangle = -\sum_\alpha c_\alpha \mid \lambda_\alpha \mid^2 - m \mid \xi \mid^2 + \sum_\alpha (\rho_\alpha \lambda_\alpha \xi + \overline{\rho_\alpha \lambda_\alpha \xi})$$

for some s-tuple $\rho = (\rho_1, \cdots, \rho_s)$. Now, for $\epsilon > 0$,

$$\mid (\rho_\alpha \lambda_\alpha \xi + \rho_\alpha \lambda_\alpha \xi) \mid \leqq 2\epsilon \mid \lambda_\alpha \mid^2 + \frac{1}{2\epsilon} \mid \rho_\alpha \mid^2 \mid \xi \mid^2$$

and so

$$\langle \bar\partial\partial\psi_i(m), \zeta \wedge \zeta \rangle \leqq \sum (-c_\alpha + 2\epsilon) \mid \lambda_\alpha \mid^2 + (-m + \sum \frac{\mid \rho\alpha \mid^2}{2\epsilon}) \mid \xi \mid^2.$$

If we first choose $2\epsilon < \min\limits_{1 \leqq \alpha \leqq s} c_\alpha$ and then $m > \sum \dfrac{\mid \rho\alpha \mid^2}{2\epsilon}$, it follows that $\partial\partial\psi_i(m)(x_0)$ has signature $(t, s+1)$. The Proposition follows by compactness.

COROLLARY. *If $X \subset W$ is a hypersurface with positive normal bundle, then $[X] \rightarrow W$ is positive for W sufficiently small.*

2. Continuous systems and the pseudo-characteristics of tubular neighborhoods. Let $X \subset W$ be an embedding and suppose that $\dim X = n$, $\dim W = n + r$. If the normal bundle $\boldsymbol{N} \rightarrow X$ of X in W is sufficiently positive

(§ II. 3), then the maximal continuous system $\{X_t\colon t \in \Sigma\}$ of X in W (§ III. 2) fills out W and it is geometrically reasonable to expect that, by interpreting some of the X_t as pseudo-characteristics of a suitable tubular neighborhood T of $X \subset W$, the E. E. Levi form on ∂T will have signature $(r-1, n)$. This section will be devoted to making this precise; in § VIII. 3 below we shall give an analytical proof of a more general fact.

(a) Suppose that $W \subset W'$ is a finite open complex manifold defined by $\{\phi < 0\}$ where ϕ is a smooth real function on W' and where $d\phi \neq 0$ in a neighborhood of ∂W. Let $m = \dim W$ and assume that $\partial\bar{\partial}\phi$ has everywhere $m-s$ positive eigenvalues in a neighborhood of ∂W. Denote by $\boldsymbol{T}^+ \subset \boldsymbol{T}(W)$ the bundle of positive eigenspaces for $\partial\bar{\partial}\phi$.

LEMMA 8.1. *Suppose that, at* $w_0 \in \partial W$, $\boldsymbol{T}^+_{w_0} \subset \boldsymbol{T}_{w_0}(\partial W)$. *Then there exists (locally) holomorphic functions* $\psi_1, \cdots, \psi_s$ *such that, in*

$$S = \{\psi_1 = \cdots = \psi_s = 0\},\ \boldsymbol{T}_{w_0}(S) \subset \boldsymbol{T}_{w_0}(\partial W),\ S \cap \partial W = \{w_0\},$$

and S *and* ∂W *have first order contact at* w_0.

Proof. Choose holomorphic coordinates $z^1, \cdots, z^m$ centered at w_0 and such that

$$\partial\bar{\partial}\phi(w_0) = \sum_{1 \leqq i \leqq m-s} dz^i \wedge d\bar{z}^i + \sum_{m-s < j \leqq m} \gamma_j dz^j \wedge d\bar{z}^j \qquad (\gamma_j \leqq 0).$$

Since $\boldsymbol{T}^+_{w_0} \subset \boldsymbol{T}_{w_0}(\partial W)$, we have

$$\phi(z) = \sum_{m-s<j\leqq m} \alpha_j z^j + \sum_{m-s<j\leqq m} \overline{\alpha_j z^j} + \sum \beta_{ij} z^i z^j + \overline{\sum \beta_{ij} z^i z^j}$$
$$+ \sum_{1\leqq i \leqq m-s} z^i \bar{z}^i + \sum_{m-s<j\leqq m} \gamma_j z^j \bar{z}^j + [3]$$

where $\beta_{ij} = \beta_{ji}$ and [3] means terms of order at least three. Also, we may assume that $\alpha_m \neq 0$. Now set

$$\psi_\alpha = z_{m-s+\alpha},\ 1 \leqq \alpha < s,\quad \psi_s = z_m + \frac{1}{\alpha_m} \sum_{1 \leqq i,j \leqq m-s} \beta_{ij} z^i z^j.$$

Then $\phi \mid S = \sum_{1\leqq i,j\leqq m-s} z^i \bar{z}^i + [3]$ and so $\phi(z) > 0$ for $z \in S - \{w_0\}$. Furthermore, $\boldsymbol{T}_{w_0}(S) = \boldsymbol{T}^+_{w_0} \subset \boldsymbol{T}_{w_0}(\partial W)$ and it is clear that ∂W and S have only first order contact at w_0.

(b) We call S a *local* $(m-s)$*-dimensional pseudo-characteristic* for W at w_0.

LEMMA 8.2. *Suppose that, at each* $w_0 \in \partial W$, *there is a local* $(m-s)$-*dimensional pseudo-characteristic. Then the Levi form on* ∂W *has* $m-s$ *positive eigenvalues.*

Proof. Let $S = \{\psi_1 = \cdots = \psi_s = 0\}$ be a local pseudo-characteristic at w_0. We may choose local coordinates $z^1, \cdots, z^m$ so that $\psi_\alpha = z^\alpha$ $(1 \leqq \alpha \leqq s)$. Then $\boldsymbol{T}_{w_0}(S)$ is spanned by the vectors $\xi = \sum_{s+1 \leqq \rho \leqq m} \xi^\rho \frac{\partial}{\partial z^\rho}$. Since $\boldsymbol{T}_{w_0}(S) \subset \boldsymbol{T}_{w_0}(\partial W)$,

$$\phi(z) = \sum_{1 \leqq \alpha \leqq s} (\gamma_\alpha z^\alpha + \overline{\gamma_\alpha z^\alpha}) + \sum_{i,j} (b_{ij} z^i z^j + \overline{b_{ij} z^i z^j} + \sum_{i,j} c_{i\bar{j}} z^i \bar{z}^j + [3]$$
$$(b_{ij} = b_{ji}, c_{i\bar{j}} = \bar{c}_{j\bar{i}}).$$

Thus

$$\phi \mid S = \sum_{s+1 \leqq \rho, \sigma \leqq m} (b_{\rho\sigma} z^\rho z^\sigma + \overline{b_{\rho\sigma} z^\rho z^\sigma}) + \sum_{s+1 \leqq \rho, \sigma \leqq m} c_{\rho\bar{\sigma}} z^\rho \bar{z}^\sigma + [3].$$

By a linear change of coordinates, we may assume $c_{\rho\bar{\sigma}} = \delta_\sigma{}^\rho \lambda_\rho$. Since $S \cap \partial W = \{w_0\}$, ϕ has the origin as a local maximum and so $\lambda_\rho \geqq 0$. Suppose say that $\lambda_m = 0$ and let Λ be the complex line $\{(0, \cdots, 0, \lambda)\}$. Then $\Lambda \subset S$ and $\phi \mid \Lambda = b\lambda^2 + \overline{b\lambda^2} + [3]$. We may choose λ_0 so that $b\lambda_0{}^2 + \overline{b\lambda_0{}^2} = 2 Re\, b\lambda_0{}^2 = 0$, and then, on the real line $\Lambda_0 = \{(0, \cdots, 0, t\lambda_0)\}$ $(t \in \boldsymbol{R})$, $\phi \mid \Lambda_0 = [3]$. Thus S and ∂W osculate to second order along Λ_0 which is contrary to assumption and so all $\lambda_\rho > 0$ $(\rho = s+1, \cdots, m)$. Q.E.D.

(c) Let $\boldsymbol{E} \xrightarrow{\pi} X$ be a holomorphic vector bundle with fibre $\boldsymbol{C}^s$ and assume that $\dim X = n$.

LEMMA 8.3. *Assume that* $\boldsymbol{E}$ *is weakly positive and let* $W \subset \boldsymbol{E}$ *be the unit tubular neighborhood. Then* W *satisfies the conditions of Lemma* 8.1 *and therefore admits local pseudo-characteristics of dimension* n.

The proof of this Lemma follows from § VI. 5, example (v), (6.43), and (6.44).

(d) Assume that $\boldsymbol{E} \to X$ is sufficiently positive. Then, given $w_0 \in \partial W$ (W = unit tubular neighborhood), there exists a neighborhood U of w_0 in $\boldsymbol{E}$ and $\sigma \in H^0(X, \Omega(\boldsymbol{E}))$ such that $\sigma(X) \cap U$ is a local pseudo-characteristic for W at w_0.

Proof. We require that $\boldsymbol{E}$ be positive enough that $H^1(X, \boldsymbol{m}_x{}^3 \cdot \Omega(\boldsymbol{E})) = 0$ $(x \in X)$ where $\boldsymbol{m}_x$ is the ideal sheaf of $\{x\} \subset X$. This implies that

$$(8.2) \qquad H^0(X,\Omega(\boldsymbol{E})) \xrightarrow{\rho_x} \Omega(\boldsymbol{E})_x/\boldsymbol{m}_x{}^3\cdot\Omega(\boldsymbol{E})_x \to 0 \qquad (x\in X).$$

Referring to the proof of Lemma 8.1 as applied to the case at hand, we see that S is defined by an $\boldsymbol{E}_{\pi(w_0)}$-valued quadratic polynomial in the local coordinates centered at $\pi(w_0)$; i.e. S is defined by an element $\psi\in\Omega(\boldsymbol{E})_x/\boldsymbol{m}_x{}^3\cdot\Omega(\boldsymbol{E})_x$. By (8.2), there exists $\sigma\in H^0(X,\Omega(\boldsymbol{E}))$ with $\rho_x(\sigma)=\psi$. Q.E.D.

Definition. We let $X_\sigma=\sigma(X)$ and call X_σ a *global pseudo characteristic* for W at w_0.

(e) PROPOSITION 8.2. *Let $X\subset W$ be an embedding with sufficiently positive normal bundle $\boldsymbol{N}\to X$. Then there exists a neighborhood T of X in W such that the E. E. Levi form has on ∂T signature* $(r-1,n)$.

Proof. We assume that $\boldsymbol{N}\to X$ is positive enough so that every point $z_0\in Z$ (= unit neighborhood of $X\in\boldsymbol{N}$) admits a global pseudo-characteristic $X_\sigma(\sigma\in H^0(X,\Omega(\boldsymbol{N}))$. Let $\Sigma\subset H^0(X,\Omega(\boldsymbol{N}))$ be the set $\{\sigma\in H^0(X,\Omega(\boldsymbol{N})):$ X_σ is a pseudo-characteristic at some $z_0\in\partial Z\}$. Then denote by $T_\delta\subset W$ the set which is enveloped by the varieties $\{X_\sigma\}_{\sigma\in(\delta\Sigma)}$. (By this, we mean that we consider the mappings $F\colon X\times\Sigma\to\boldsymbol{N}$, $G\colon X\times\Sigma\to W$ and set $T_\delta=G(F^{-1}(Z_\delta))$.) Then T_δ is a neighborhood of X in W and an easy approximation argument (cf. § VIII.3 below) shows that the varieties X_σ ($\sigma\in(\delta\Sigma)$) will, for δ small enough, give a set of global pseudo-characteristics for T_δ.

3. Another proof of the concavity, convexity of tubular neighborhoods. We follow the notations of § VIII.1 above. Let $\sigma\in H^0(W,\Omega(\boldsymbol{L}))$ be the canonical section given by $\sigma\mid U_i=f_i$ and, for $\tau\in\boldsymbol{R}^+$, we set $\sigma_\tau=\tau\sigma$. Then $X\subset W$ is given by $\{w\in W:\sigma(w)=0\}$. Assume that we are given a metric γ in $\boldsymbol{N}\to X$ whose curvature has signature (s,t) $(s+t=n=\dim X)$. Then γ defines in $\boldsymbol{N}$ a real non-negative function ϕ such that $\partial\bar\partial\phi$ has signature $(s,t+1)$ on $\boldsymbol{N}\to X$ (cf. § VI.5, example (v)). By a partition of unity, we extend ϕ to a metric function ψ defined on $\boldsymbol{L}\supset\boldsymbol{N}$ (cf. § VIII.1) and we define ϕ_τ on W by $\phi_\tau=\psi\circ\sigma_\tau$.

PROPOSITION 8.3. *There exists constants τ, c_1, c_2 such that:* (i) $0<c_1<c_2$ *and the locus $\{w\in W:\phi_\tau(w)<c_2\}$ defines an open submanifold $W'\subset W$; and* (ii) *if $W''=\{w:\phi_\tau(w)<c_1\}$, then on $W'-W''$, $\partial\bar\partial\phi_\tau$ has signature* $(t+1,s)$.

Proof. We choose a covering of $\boldsymbol{L}$ by neighborhoods $V_i = \pi^{-1}(U_i)$ ($\pi: \boldsymbol{L} \to W$ being the projection). If (z_i, w_i) $(z_i = (z_i^1, \cdots, z_i^n)$ and $X \cap U_i = \{w_i = 0\})$ are coordinates in U_i then in V_i we have coordinates (z_i, w_i, ξ_i) with $\xi_i = (w_i/w_j)\xi_j$ in $V_i \cap V_j$. On V_i we write

$$\psi(z_i, w_i, \xi_i) = \phi(z_i, \xi_i) + \rho(z_i, w_i)\eta(z_i, w_i, \xi_i)$$

where ρ is a non-negative function on W which defines the distance to X. Since ψ comes from a metric, it is locally divisible by $|\xi_i|^2$ and we have

$$\psi(z_i, w_i, \xi_i) = (\phi_i(z_i) + \rho(z_i, w_i)\eta_i(z_i, w_i))\,|\xi_i|^2.$$

Now $\sigma_\tau(W) \cap V_i$ is given by $\xi_i - \tau w_i = 0$ and thus

$$\phi_\tau \mid U_i = \{\phi_i(z_i) + \rho(z_i, w_i)\eta_i(z_i, w_i)\}\tau^2\,|w_i|^2.$$

Then

$$\begin{aligned}\bar{\partial}\partial\phi_\tau = (\bar{\partial}\partial\phi_i)\tau^2\,|w_i|^2 - \tau^2\phi_i\, dw_i \wedge d\bar{w}_i \\ + \tau^2\,|w_i|^2\{\bar{\partial}\partial\eta_i\rho + \eta_i\bar{\partial}\partial\rho + \bar{\partial}r \wedge \partial\eta_i - \partial r \wedge \bar{\partial}\eta_i\}.\end{aligned}$$

Since ρ, $\partial\rho$, $\bar{\partial}\rho$, $\bar{\partial}\partial\rho$ are all zero along X, we may find a constant $c > 0$ such that, if $W_c = \{w: \phi_1(w) < c\}$, then there is at every point $w \in W_c$ an n-dimensional subspace $\boldsymbol{S}_w \subset \boldsymbol{T}_w(W)$ on which

$$(\bar{\partial}\partial\phi_i + \{\bar{\partial}\partial\eta_i\rho + \eta_i\bar{\partial}\partial\rho + \bar{\partial}\rho \wedge \partial\eta_i - \partial\rho \wedge \bar{\partial}\eta_i\})$$

will have signature (s, t). Furthermore, by taking if necessary a smaller c, $\dfrac{\partial}{\partial w_i}$ will be normal to $\boldsymbol{S}_w$. Then the same argument as in § VIII.1 shows that we may find a large $\tau >> 0$ and $c' < c$ such that $\bar{\partial}\partial\phi_\tau$ has in $W_c - W_{c'}$ the signature $(s, t+1)$. Q.E.D.

COROLLARY (Grauert [6]). *If $\boldsymbol{N} \to X$ is negative, then there exists a strongly pseudo-convex neighborhood of X in W.*

4. Proof of Theorem VII. We first choose a metric in $\boldsymbol{L}^* \xrightarrow{\pi} W$ whose curvature has signature $(t, s+1)$ according to Proposition 8.1. By Proposition 6.4 we may choose Hermitian metrics on $W' \supset W$ such that $\Theta^{\#}(\eta, \eta)$ is negative definite for $\eta \in A^q(\boldsymbol{L}^*)$, $0 \leqq q < s$ and $\Theta(\eta, \eta)$ is positive definite for $\eta \in A^q(\boldsymbol{L}^*)$, $n + 1 - t < q \leqq n + 1$.

Now let $W \supset X$ be chosen sufficiently small and let ϕ_τ be as in § VIII.3.

Then, by the explicit construction, we see that for all $w \in W - X$ there exists an n-dimensional subspace $\boldsymbol{S}_w \subset \boldsymbol{T}_w(W)$ on which Θ and $\partial\bar{\partial}\phi_\tau$ each have signature (t, s). Furthermore we may choose subspaces $\boldsymbol{S}_w^+$, $\boldsymbol{S}_w^-$ such that $\boldsymbol{S}_w^+ + \boldsymbol{S}_w^- = \boldsymbol{S}_w$ and such that Θ, $\partial\bar{\partial}\phi_\tau$ are on $\boldsymbol{S}^+$ positive definite and are negative definite on $\boldsymbol{S}^-$. Finally, for W still sufficiently small, $\boldsymbol{S}_w$ will be close to $\boldsymbol{T}_w(\partial W)$. It then follows that, with the chosen Hermitian metrics, $L^{\#}(\eta, \eta)$ will be negative definite for $\eta \in \boldsymbol{D}^q(\boldsymbol{L}^*)$, $0 \leqq q < s$ and $L(\eta, \eta)$ will be positive definite for $\eta \in \boldsymbol{D}^q(\boldsymbol{L}^*)$, $n - t < q$ (we use here the boundary conditions as expressed by (7.20)). Finally, without effecting any of our choices, we may multiply the metric in $\boldsymbol{L}^*$ by $\bar{e}^{\lambda\phi_\tau}$ $(\lambda >> 0)$ so that the new metric curvature will have signature $(t + 1, s)$ and $\Theta(\eta, \eta)$ will be positive definite for $\eta \in A^q(\boldsymbol{L}^*)$, $n - t < q$. This proves Theorem VII.

REFERENCES.

[1] A. Andreotti, "Théorèmes de dépendance algébrique sur les espaces complexes pseudo-concaves," *Bulletin de la Sociètè Mathématique de France*, vol. 91 (1963), pp. 1-38.

[2] A. Andreotti and E. Vesentini, "Carleman stimates for the Laplace-Beltrami operator on complex manifolds," to appear.

[3] A. Andreotti and H. Grauert, "Théorèmes de Finitude pour la cohomologie des esplaces complexes," *Bulletin de la Sociètè Mathématique de France*, vol. 90 (1962), pp. 193-259.

[4] S. Bochner and K. Yano, "Curvature and Betti numbers," *Annals of Mathematics Studies*, No. 32 (1953).

[5] S. S. Chern, "Characteristic classes of Hermitian manifolds," *Annals of Mathematics*, vol. 47 (1946), pp. 85-121.

[6] H. Grauert, "Über modifikationen und exzeptionelle analytische mengen," *Mathematische Annalen*, vol. 146 (1962), pp. 331-368.

[7] P. Griffiths, "The extension problem in complex analysis, I," to appear in Proceedings Conference on Complex Analysis, University of Minnesota (1964).

[8] ———, "Hermitian differential geometry and the theory of positive and ample holomorphic bundles," to appear in *Journal of Mathematics and Mechanics.*

[9] F. Hirzebruch, "Neue topologische methoden in der algebraischen geometrie," *Ergebnisse der Mathematik*, vol. 9 (1956), Springer.

[10] W. Hodge and D. Pedoe, *Methods of algebraic geometry* (III), Cambridge University Press (1952).

[11] C. C. Hsuing, "Curvature and homology of Riemannian manifolds with boundary," *Mathematische Zeitschrift*, vol. 82 (1963), pp. 67-82.

[12] K. Kodaira, "A theorem of completeness of characteristics systems for analytic families of compact submanifolds of complex manifolds," *Annals of Mathematics*, vol. 75 (1962), pp. 146-162.

[13] ———, "On Kähler varieties of restricted type," *Annals of Mathematics*, vol. 60 (1954), pp. 28-48.

[14] K. Kodaira and D. C. Spencer, "Groups of complex line bundles over compact Kähler varieties," *Proceedings of the National Academy of Sciences, U.S.A.*, vol. 39 (1953), pp. 1273-1278.

[15] J. J. Kohn, "Harmonic integrals on strongly pseudo-convex manifolds I, II," *Annals of Mathematics*, vol. 78 (1963), pp. 112-148.

[16] Y. Nakai, "A criterion of an ample sheaf on a projective scheme," *American Journal of Mathematics*, vol. 85 (1963), pp. 14-26.

[17] L. Nirenberg and D. C. Spencer, "On rigidity of holomorphic embeddings," Conference on Function Theory, Tata Institute (1962), pp. 133-137.

[18] H. Rossi, to appear.

A NOTE ON COVERINGS OVER IMPERFECT GROUND FIELDS.

By Peter L. Falb.*

1. **Introduction.** Let K be a function field in r-variables over a ground field k of characteristic $p \neq 0$. We assume that K is a regular extension of k and that K^* is a finite separable algebraic extension of K. We wish to investigate several properties of the extension K^*/K using the techniques of [1] and [2]. In particular, we shall

(a) show that $p_g(K/k) \leqq p_g(K^*/k)$ where p_g denotes the geometric genus ([2]);

(b) prove a criterion for ramification ([3]); and

(c) prove the purity of the branch locus under the assumption that $[k : k^p] = p^\rho < \infty$.

We recall ([2]) that if $\mathfrak{p}$ is a prime divisor of K/k and if $u_1, u_2, \cdots, u_\rho$ are p-parameters for $\mathfrak{p}$ with k_1 as field of p-constants, then we can speak of uniformizing coordinates of $\mathfrak{p}$ relative to the p-parameters $u_1, u_2, \cdots, u_\rho$ and the field of p-constants k_1, i. e., we can consider elements $\xi_1, \xi_2, \cdots, \xi_{\rho+r}$ of the valuation ring $\mathfrak{O}$ of $\mathfrak{p}$ for which there exist derivations D_i of K such that $D_i\xi_j = \delta_{ij}$ $(i, j = 1, \cdots, \rho + r)$ and $D_i\mathfrak{O} \subset \mathfrak{O}$. Uniformizing coordinates $\xi_1, \cdots, \xi_{\rho+r}$ of $\mathfrak{p}$ are characterized by the conditions:

(i) $k(\mathfrak{p})$ is separable over $k_1(\bar{\xi}_1, \cdots, \bar{\xi}_{\rho+r})$ where bar denotes $\mathfrak{M}$-residue, $\mathfrak{M}$ being the maximal ideal of $\mathfrak{O}$; and

(ii) $k_1[\xi_1, \cdots, \xi_{\rho+r}]$ contains a set of uniformizing parameters of $\mathfrak{p}$.

We recall also that if $x_1, \cdots, x_r$ is a separating transcendence basis of K/k, if $\xi_1, \cdots, \xi_{\rho+r}$ are uniformizing coordinates of $\mathfrak{p}$, and if $\omega = A\,dx_1 \cdots dx_r$ is an r-fold differential of K/k, then the order of ω at $\mathfrak{p}$, $v_\mathfrak{p}(\omega)$, is the integer $v_\mathfrak{p}(A) + v_\mathfrak{p}(dx_1 \cdots dx_r)$ where $v_\mathfrak{p}(dx_1 \cdots dx_r) = v_\mathfrak{p}(B)$ where

$$dx_1 \cdots dx_r du_1 \cdots du_\rho = B\,d\xi_1 \cdots d\xi_{\rho+r}$$

[the u_i being our p-parameters]. We note finally that if $t_1, \cdots, t_{\rho+r}$ are

---

Received February 9, 1965.

[*] Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts. Operated with support from the U. S. Air Force.

elements of $\mathfrak{O}$, then the $\rho + r$-fold differential $dt_1 \cdots dt_{\rho+r}$ of K/k_1 is regular at $\mathfrak{p}$ in the sense that $v_{\mathfrak{p}}(dt_1 \cdots dt_{\rho+r}) \geqq 0$.

2. Behaviour of the geometric genus. We shall now investigate the behaviour of the geometric genus under the extension from K to K^*. First of all, we observe that, in view of the main result of [2], we may assume that k is maximally algebraic in K^*, since, if this were not so, then, letting k^* be the algebraic closure of k in K^*, we could consider the tower $K \subset Kk^* \subset K^*$ and apply the main result of [2] to Kk^*/K. Secondly, we have:

LEMMA 1. *Let $\mathfrak{p}^*$ be a prime divisor of K^*/k and let $\mathfrak{p}$ be its contraction to K. Let $x_1, \cdots, x_r$ be a separating transcendence basis of K/k (a fortiori of K^*/k) such that $\mathfrak{p}$, and hence $\mathfrak{p}^*$, is of the first kind with respect to $k[x_1, \cdots, x_r]$. If R is the integral closure of $k[x_1, \cdots, x_r]$ in K and R^* is the integral closure of R in K^*, then*

$$v_{\mathfrak{p}^*}(d^*x_1 \cdots d^*x_r) = e_{\mathfrak{p}^*/\mathfrak{p}} v_{\mathfrak{p}}(dx_1 \cdots dx_r) + v_{\mathfrak{p}^*}(\mathfrak{D}_{R^*/R}) \tag{1}$$

where $\mathfrak{D}_{R^*/R}$ is the different of R^* over R and the asterisk superscript above a differential indicates that the differential is one of K^*/k.[1]

Proof. We know from Proposition 1 of [2] that

$$v_{\mathfrak{p}^*}(\mathfrak{D}_{R^*/k[x_1,\cdots,x_r]}) = v_{\mathfrak{p}^*}(d^*x_1 \cdots d^*x_r) \tag{2}$$

and that

$$v_{\mathfrak{p}}(dx_1 \cdots dx_r) = v_{\mathfrak{p}}(\mathfrak{D}_{R/k[x_1,\cdots,x_r]}). \tag{3}$$

However,

$$v_{\mathfrak{p}^*}(\mathfrak{D}_{R^*/k[x_1,\cdots,x_r]}) = v_{\mathfrak{p}^*}(\mathfrak{D}_{R^*/R}) + e_{\mathfrak{p}^*/\mathfrak{p}} v_{\mathfrak{p}}(\mathfrak{D}_{R/k[x_1,\cdots,x_r]}) \tag{4}$$

([4a], p. 299 and p. 309) and the lemma is established.

Let $\Omega = \{\omega : \omega$ is an r-fold differential of K/k and ω is regular at all prime divisors of $K/k\}$ and let $\Omega^* = \{\omega^* : \omega^*$ is an r-fold differential of K^*/k and ω^* is regular at all prime divisors of $K^*/k\}$. Then $p_g(K/k) = \dim_k \Omega$ and $p_g(K^*/k) = \dim_k \Omega^*$.

THEOREM 1. *$p_g(K/k) \leqq p_g(K^*/k)$.*

Proof. Suppose that $\omega_1, \cdots, \omega_s$ are k-independent elements of Ω and let ω_i^* be the cotrace of ω_i taken from K to K^*. Clearly, $\omega_1^*, \cdots, \omega_s^*$ are k-independent and so it will suffice to show that $v_{\mathfrak{p}^*}(\omega_i^*) \geqq 0$ for any prime divisor of K^*/k. Let $\mathfrak{p}^*$ be a prime divisor of K^*/k and let $\mathfrak{p}$, R, R^*, and $x_1, \cdots, x_r$ be as in Lemma 1. Then, by Lemma 1,

---

[1] We use this notation throughout the sequel.

$$v_{\mathfrak{p}^*}(\omega_i^*) = e_{\mathfrak{p}^*/\mathfrak{p}} v_{\mathfrak{p}}(\omega_i) + v_{\mathfrak{p}^*}(\mathfrak{D}_{R^*/R}). \tag{5}$$

But, $v_{\mathfrak{p}}(\omega_i) \geqq 0$ and $v_{\mathfrak{p}^*}(\mathfrak{D}_{R^*/R}) \geqq 0$. Thus, the theorem is established.

COROLLARY 1. *If V is a normal model of K/k and if V^* is a normalization of V in K^*, then $p_g(V/k) \leqq p_g(V^*/k)$ [see [2] for a definition of $p_g(V/k)$].*

3. A criterion for ramification ([3]). We extend the well-known criterion for ramification of [3] to the situation at hand. Since we deal with prime divisors of the first kind on normal models and since prime divisors do not ramify under finite separable algebraic constant field extensions, we may assume that k is maximally algebraic in K^*. We let V be a normal model of K/k, V^* be a normalization of V in K^*, W be an irreducible, $r-1$ dimensional k-subvariety of V, and W^* be an irreducible, $r-1$ dimensional k-subvariety of V^* lying over W. We say that W^* *is unramified over* W if the local ring $\mathfrak{O}(W^*, V^*)$ of W^* on V^* is unramified over the local ring $\mathfrak{O}(W, V)$ of W on V. We suppose that p-parameters $u_1, \cdots, u_\rho$ and a field k_1 of p-constants for V, V^*, W, and W^* have been fixed ([1], Section 1).

THEOREM 2. *Let $\xi_1, \cdots, \xi_{\rho+r}$ be uniformizing coordinates of W on V relative to the u_i and k_1 ([1], Proposition 1). Then W^* is ramified over W if and only if*

$$v_{W^*}(d^*\xi_1 \cdots d^*\xi_{\rho+r}) > 0 \tag{6}$$

where v_{W^*} is the valuation of the prime divisor determined by W^*.

Proof. If W^* is unramified over W, then, by the criterion for uniformizing coordinates ([1], Theorem 1), $\xi_1, \cdots, \xi_{\rho+r}$ are uniformizing coordinates of W^* on V^* and so, $v_{W^*}(d^*\xi_1 \cdots d^*\xi_{\rho+r}) = 0$.

If W^* is ramified over W, there are two cases, namely:

(a) $e_{W^*/W} > 1$ or

(b) $e_{W^*/W} = 1$ and $[k(W^*) : k(W)]_i > 1$.

Case (a). We may assume that ξ_1 (say) is a uniformizing parameter of W on V ([1], Corollary, p. 549). Let ξ_1^* be a uniformizing parameter of W^* on V^* and let $e = e_{W^*/W}$. Then $\xi_1 = u\xi_1^{*e}$ where u is a unit. It follows that

$$d^*\xi_1 \cdots d^*\xi_{\rho+r} = \xi_1{}^{*e-1}\{eud\xi_1{}^*d^*\xi_2 \cdots d^*\xi_{\rho+r} + \xi_1{}^*dud^*\xi_2 \cdots d^*\xi_{\rho+r}\} \tag{7}$$

and hence, that

$$\begin{aligned} &v_{W^*}(d^*\xi_1 \cdots d^*\xi_{\rho+r}) \\ &\qquad = e - 1 + v_{W^*}(eud\xi_1{}^*d^*\xi_2 \cdots d^*\xi_{\rho+r} + \xi_1{}^*dud^*\xi_2 \cdots d^*\xi_{\rho+r}). \end{aligned} \tag{8}$$

But, $e > 1$ and $\xi_1{}^*, \xi_2, \cdots, \xi_{\rho+r}$, u are in $\mathfrak{O}(W^*, V^*)$ together imply that $v_{W^*}(d^*\xi_1 \cdots d^*\xi_{\rho+r}) > 0$.

Case (b). We again assume that ξ_1 is a uniformizing parameter of W on V. Then ξ_1 is a uniformizing parameter of W^* on V^* and there are elements $\eta_2, \cdots, \eta_{\rho+r}$ in $\mathfrak{O}(W^*, V^*)$ such that $\xi_1, \eta_2, \cdots, \eta_{\rho+r}$ are uniformizing coordinates of W^* on V^* ([1], Corollary, p. 549). Letting a bar denote the $\mathfrak{M}(W^*, V^*)$ residue of an element of $\mathfrak{O}(W^*, V^*)$, we can assert that $k(W^*)$ is separable algebraic over $k_1(\bar{\eta}_2, \cdots, \bar{\eta}_{\rho+r})$. It follows that there are $\rho + r - 1$ polynomials Ψ_ν with coefficients in k_1 such that

$$\Psi_\nu(\bar{\xi}_2, \cdots, \bar{\xi}_{\rho+r}; \bar{\eta}_2, \cdots, \bar{\eta}_{\rho+r}) = 0, \quad \nu = 1, \cdots, \rho + r - 1 \tag{9a}$$

$$\det \| \delta\Psi_\nu / \delta\bar{\xi}_i \| \neq 0 \tag{9b}$$

$$\det \| \delta\Psi_\nu / \delta\bar{\eta}_j \| = 0 \tag{9c}$$

((9c) is a consequence of the fact that $k_1(\bar{\eta}_2, \cdots, \bar{\eta}_{\rho+r})$ is *not* separable algebraic over $k_1(\bar{\xi}_2, \cdots, \bar{\xi}_{\rho+r})$). We deduce that

$$\det \| \delta\Psi_\nu / \delta\xi_i \| \, d^*\xi_1 \cdots d^*\xi_{\rho+r} = a\xi_1 d^*\xi_1 d\eta_2 \cdots d\eta_{\rho+r} \tag{10}$$

where a is an element of $\mathfrak{O}(W^*, V^*)$. Since $v_{W^*}(\det \| \delta\Psi_\nu / \delta\xi_i \|) = 0$, we see immediately that $v_{W^*}(d^*\xi_1 \cdots d^*\xi_{\rho+r}) > 0$ and the proof of the theorem is complete.

4. Purity of the branch locus. We use Theorem 2 to prove the purity of the branch locus ([5]) in this section. We again let V be a normal model of K/k and V^* be a normalization of V in K^*. Since we consider finite sets of k-subvarieties of V and V^*, we (tacitly) assume that p-parameters $u_1, \cdots, u_\rho$ and a field k_1 of p-constants of V and V^* are p-parameters and field of p-constants respectively for any of the varieties we consider. (cf. [1], Section 1.)

Let W be an s-dimensional irreducible k-subvariety of V and let $S(W/V) = \{\mathfrak{p} : \mathfrak{p}$ is a prime divisor of K/k which is of the first kind on V and $\mathfrak{O}_{\mathfrak{p}}{}^2 \supset \mathfrak{O}(W, V)\}$. Then we say that a *differential* ω *of* K/k (or k_1) *is*

---

[2] $\mathfrak{O}_{\mathfrak{p}}$ is the valuation ring of $\mathfrak{p}$.

regular on W if $v_{\mathfrak{p}}(\omega) \geqq 0$ for every $\mathfrak{p}$ in $S(W/V)$. We let $\Omega(W) = \{\omega$ is a $\rho + r$-fold differential of K/k_1 which is regular on $W\}$ and we then have:

LEMMA 2. *If $\eta_1, \cdots, \eta_{\rho+r}$ are elements of $\mathfrak{O} = \mathfrak{O}(W, V)$ such that K is separable algebraic over $k_1(\xi_1, \cdots, \xi_{\rho+r})$ and if $\Omega(W)$ is a free $\mathfrak{O}$-module with base $d\eta_1 \cdot \cdots \cdot d\eta_{\rho+r}$, then $\eta_1, \cdots, \eta_{\rho+r}$ are uniformizing coordinates of W on V. Conversely, if $\xi_1, \cdots, \xi_{\rho+r}$ are uniformizing coordinates of W on V, then $\Omega(W)$ is a free $\mathfrak{O}$-module with base $d\xi_1 \cdot \cdots \cdot d\xi_{\rho+r}$.*

Proof. Since $K/k_1(\eta_1, \cdots, \eta_{\rho+r})$ is separable algebraic and the dimension of the space of derivations of K/k_1 is $\rho + r$, there are derivations $E_1, \cdots, E_{\rho+r}$ of K/k_1 such that $E_i\eta_j = \delta_{ij}$ $(i, j = 1, \cdots, \rho + r)$. We claim that $E_i\mathfrak{O} \subset \mathfrak{O}$ for $i = 1, \cdots, \rho + r$ which will show that the η_j are uniformizing coordinates of W on V. Suppose that E_1, say, did *not* map $\mathfrak{O}$ into $\mathfrak{O}$; then there would be an a in $\mathfrak{O}$ such that $E_1 a$ was not in $\mathfrak{O}$. However, $da d\eta_2 \cdot \cdots d\eta_{\rho+r} = (E_1 a) d\eta_1 \cdot \cdots d\eta_{\rho+r}$ is in $\Omega(W)$ (as $a, \eta_2, \cdots, \eta_{\rho+r}$ are in $\mathfrak{O}$ and a fortiori in $\mathfrak{O}_{\mathfrak{p}}$ for every $\mathfrak{p}$ in $S(W/V)$) but would *not* be in the free $\mathfrak{O}$-module with base $d\eta_1 \cdot \cdots d\eta_{\rho+r}$. This contradiction establishes the first assertion of the lemma.

On the other hand, suppose that $\xi_1, \cdots, \xi_{\rho+r}$ are uniformizing coordinates of W on V and let $D_1, \cdots, D_{\rho+r}$ be the derivations of K/k_1 such that $D_i\xi_j = \delta_{ij}$ $(i, j = 1, \cdots, \rho + r)$. Since V is normal, we know that if $\mathfrak{p}$ is in $S(W/V)$, then $\mathfrak{O}_{\mathfrak{p}} = \mathfrak{O}_M$ where M is a suitable multiplicative set. It follows that $D_i\mathfrak{O} \subset \mathfrak{O}$ implies that $D_i\mathfrak{O}_{\mathfrak{p}} \subset \mathfrak{O}_{\mathfrak{p}}$ for every $\mathfrak{p}$ in $S(W/V)$. In other words, $\xi_1, \cdots, \xi_{\rho+r}$ are uniformizing coordinates at every $\mathfrak{p}$ in $S(W/V)$. Now let ω be an element of $\Omega(W)$. Then $\omega = A d\xi_1 \cdot \cdots d\xi_{\rho+r}$ with A in $\mathfrak{O}_{\mathfrak{p}}$ for every $\mathfrak{p}$ in $S(W/V)$. But V is normal so that $\mathfrak{O} = \cap \mathfrak{O}_{\mathfrak{p}}$, $\mathfrak{p} \in S(W/V)$ and therefore, A is in $\mathfrak{O}$. The proof of the lemma is now complete.

LEMMA 3. *Suppose that uniformizing coordinates of W on V do not exist, that $\eta_1, \cdots, \eta_{\rho+r}$ are elements of $\mathfrak{O} = \mathfrak{O}(W, V)$ such that $K/k_1(\eta_1, \cdots, \eta_{\rho+r})$ is separable algebraic, and that ω is an element of $\Omega(W)$ which can be written in the form $a d\eta_1 \cdot \cdots d\eta_{\rho+r}$ with a in $\mathfrak{O}$. Then there is a prime divisor $\mathfrak{p}$ whose center Γ on V contains W such that $v_{\mathfrak{p}}(\omega) > 0$.*

Proof. Let $E_1, \cdots, E_{\rho+r}$ be the derivations of K/k_1 such that $E_i\eta_j = \delta_{ij}$ $(i, j = 1, \cdots, \rho + r)$. Then we may suppose that (say) $E_1\mathfrak{O} \not\subset \mathfrak{O}$ since uniformizing coordinates of W on V do not exist. Let b be an element of $\mathfrak{O}$ such that $E_1 b$ is not in $\mathfrak{O}$. Since V is normal, there is a prime divisor $\mathfrak{p}$ whose center Γ on V contains W such that $v_{\mathfrak{p}}((E_1 b)^{-1}) > 0$. However, $\omega = (E_1 b)^{-1} a db d\eta_2 \cdot \cdots d\eta_{\rho+r}$ which implies that

$$v_{\mathfrak{p}}(\omega) = v_{\mathfrak{p}}((E_1 b)^{-1}) + v_{\mathfrak{p}}(adb\,d\eta_2 \cdots d\eta_{\rho+r}) > 0$$

since $a, b, \eta_2, \cdots, \eta_{\rho+r}$ are in $\mathfrak{O}$ and $\mathfrak{O} \subset \mathfrak{O}_{\mathfrak{p}}$.

We suppose now that k is maximally algebraic in K^* and that $[k : k^p] = p^\rho < \infty$. We assume that our p-parameters are a p-basis of k/k^p and that k^p is our field of p-constants. We recall ([5]) that if W^* is an irreducible k-subvariety of V^* which lies over the irreducible k-subvariety W of V, then W^* is *unramified over* W if $\mathfrak{M}\mathfrak{O}^* = \mathfrak{M}^*$ and $k(W^*)$ is a separable extension of $k(W)$ where $\mathfrak{O}^* = \mathfrak{O}(W^*, V^*)$, $\mathfrak{M} = \mathfrak{M}(W, V)$, and $\mathfrak{M}^* = \mathfrak{M}(W^*, V^*)$. If Q^* is a point of V^* which lies over a point Q of V, then Q^* is *ramified over* Q if the locus of Q^* over k is not unramified over the locus of Q over k. We recall also that the branch locus of V^* over V is the set of all ramified points of V^* and is an algebraic set defined over k ([5]). We now have:

THEOREM 3. *If W is simple on V and if W^* is ramified over W, then there is an irreducible $r-1$ dimensional subvariety Γ^* of V^*, containing W^*, which is ramified over the corresponding subvariety Γ of V.*

Proof. Since the branch locus is an algebraic set defined over k, we may assume that $\dim W/k = 0$.[3] We consider two cases, namely:

(a) there do *not* exist uniformizing coordinates of W^* on V^*; and

(b) there exist uniformizing coordinates of W^* on V^*.

Case (a). Since W is simple on V, there are uniformizing coordinates $\xi_1, \cdots, \xi_{\rho+r}$ of W on V and the differential $d^*\xi_1 \cdots d^*\xi_{\rho+r}$ is in $\Omega(W^*)$. It follows from Lemma 3 that there is an irreducible $r-1$ dimensional subvariety Γ^* of V^* containing W^* such that $v_{\Gamma^*}(d^*\xi_1 \cdots d^*\xi_{\rho+r}) > 0$ (where v_{Γ^*} is the valuation of the prime divisor determined by Γ^*). In view of Theorem 2, Theorem 3 is established in this case.

Case (b). Let $\xi_1, \cdots, \xi_{\rho+r}$ be uniformizing coordinates of W on V; let $\omega = d\xi_1 \cdots d\xi_{\rho+r}$; and let $\eta_1, \cdots, \eta_{\rho+r}$ be uniformizing coordinates of W^* on V^*. Then $\omega^* = d^*\xi_1 \cdots d^*\xi_{\rho+r}$ is in $\Omega(W^*)$ and $\omega^* = a\,d\eta_1 \cdots d\eta_{\rho+r}$ with a in $\mathfrak{O}^* = \mathfrak{O}(W^*, V^*)$. If a is in $\mathfrak{M}^* = \mathfrak{M}(W^*, V^*)$, then there is an irreducible $r-1$ dimensional subvariety Γ^* of V^* containing W^* such that $v_{\Gamma^*}(a) > 0$ (as V^* is normal). It follows that $v_{\Gamma^*}(\omega^*) > 0$ and thus, that Γ^* is ramified over Γ (Theorem 2). In other words, it will suffice to show that a is in $\mathfrak{M}^*$.

Suppose that a were not in $\mathfrak{M}^*$. Then, in view of Lemma 2, ω^* would be an $\mathfrak{O}^*$ base of Ω^* and consequently, $\xi_1, \cdots, \xi_{\rho+r}$ would be uniformizing coordinates of W^* on V^*. We claim that it would follow that W^* was unramified over W contradicting the hypothesis of the theorem. To verify this claim, we apply Remark 3 of [5] (p. 795). We let δ^* be a local W^* derivation which vanishes on $\mathfrak{O} = \mathfrak{O}(W, V)$ and we let α be any element of $\mathfrak{O}^*$. Since K^* is separable algebraic over $k^p(\xi_1, \cdots, \xi_{\rho+r})$, $\{\xi_1, \cdots, \xi_{\rho+r}\}$ is a p-basis of K^*/K^{*p}. It follows that

---

[3] We make this assumption to facilitate the use of [5], Remark 3, p. 795.

$$\alpha = \sum_{\langle i \rangle} C_{\langle i \rangle}{}^p \xi_1{}^{i_1} \cdots \xi_{\rho+r}{}^{i_{\rho+r}} \tag{11}$$

where $\langle i \rangle = (i_1, \cdots, i_{\rho+r})$, $i_1 < \cdots < i_{\rho+r}$, and $i_j \in \{0, \cdots, p-1\}$. Since $\mathfrak{O}^*$ is integrally closed and the derivations $D_i{}^*$ of K^*/k^p such that $D_i{}^* \xi_j = \delta_{ij}$ $(i, j = 1, \cdots, \rho + r)$ map $\mathfrak{O}^*$ into $\mathfrak{O}^*$ (as the ξ_j are also uniformizing coordinates of W^*), we deduce that $C_{\langle i \rangle}$ is in $\mathfrak{O}^*$. But, $\delta^*(\xi_i) = 0$ for $i = 1, \cdots, \rho + r$ and

$$\delta^*(C_{\langle i \rangle}{}^p) = p\bar{C}_{\langle i \rangle}{}^{p-1} \delta^*(C_{\langle i \rangle}) = 0 \tag{12}$$

imply that $\delta^*(\alpha) = 0$. As α is any element of $\mathfrak{O}^*$, our claim is established in view of Remark 3 of [5]. Thus, the proof of the theorem is complete.

REFERENCES.

[1] O. Zariski and P. Falb, "On differentials in function fields," *American Journal of Mathematics*, vol. 83 (1961), pp. 542-556.

[2] P. Falb, "On the behaviour of the geometric genus under ground field extension," *American Journal of Mathematics*, vol. 84 (1962), pp. 421-432.

[3] Y. Nakai, "Ramifications, differentials and differente on algebraic varieties of higher dimensions," *Memoirs of the College of Science*, Kyoto, Series A, Mathematics, vol. 32 (1960), pp. 391-411.

[4a] O. Zariski and P. Samuel, *Commutative Algebra*, Vol. I; D. Van Nostrand Co., Princeton, 1958.

[4b] ———, *Commutative Algebra*, Vol. II; D. Van Nostrand Co., Princeton, 1960.

[5] O. Zariski, "On the purity of the branch locus of algebraic functions," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 44 (1958), pp. 791-796.

ON THE THEOREMS OF WATSON AND DRAGONETTE FOR RAMANUJAN'S MOCK THETA FUNCTIONS.

By George E. Andrews.[1]

I. Introduction. The mock theta functions were first studied by Ramanujan; G. H. Hardy in [9; p. 354] defines such functions as follows

> . . . a "mock θ-function" is a function defined by a q-series convergent for $|q|<1$, for which we can calculate asymptotic formulae, when q tends to a "rational point" $e^{2r\pi i/s}$, of the same degree of precision as those furnished, for the ordinary θ-functions, by the theory of linear transformation.

The mock theta functions we shall study are the third order functions given below:

$$f(q)=\sum_{n=0}^{\infty}\frac{q^{n^2}}{(1+q)^2(1+q^2)^2\cdots(1+q^n)^2}$$

$$\phi(q)=\sum_{n=0}^{\infty}\frac{q^{n^2}}{(1+q^2)(1+q^4)\cdots(1+q^{2n})}$$

$$\psi(q)=\sum_{n=0}^{\infty}\frac{q^{n^2}}{(1-q)(1-q^3)\cdots(1-q^{2n-1})}$$

$$\omega(q)=\sum_{n=0}^{\infty}\frac{q^{2n(n+1)}}{(1-q)^2(1-q^3)^2\cdots(1-q^{2n+1})^2}$$

$$v(q)=\sum_{n=0}^{\infty}\frac{q^{n(n+1)}}{(1+q)(1+q^3)\cdots(1+q^{2n+1})}$$

There are two other third order functions $\chi(q)$ and $\rho(q)$ [12; p. 62]. They satisfy the following equations [12; p. 63]

$$\chi(q)=\tfrac{1}{4}f(q)+\tfrac{3}{4}\theta_4{}^2(0,q^3)\prod_{r=1}^{\infty}(1-q^r)^{-1}$$

$$\rho(q)=-\tfrac{1}{2}\omega(q)+\tfrac{3}{2}[\tfrac{1}{2}q^{-3/8}\theta_2(0,q^{3/2})]^2\prod_{r=1}^{\infty}(1-q^{2r})^{-1}.$$

Received November 16, 1964.

[1] Presented to the faculty of the Graduate School of the University of Pennsylvania in partial fulfillment of the requirements for the degree of Doctor of Philosophy, 1964.

This research was partially completed under a National Science Foundation Graduate Fellowship and partially under NSF grant GP1640.

We shall not study them; however, their behavior can be derived from these equations once the behavior of f and ω is known.

In the paper [2], Dragonette derived general linear transformation formulae for the third order mock theta functions and used her results to obtain very precise asymptotic formulae for the power series coefficients of these functions. G. N. Watson in [12] had already studied the third order mock theta functions under the specific transformations; $S\colon \tau \rightarrow \tau + 1$ and $T\colon \tau \rightarrow 1/\tau$. Since these transformations generate the modular group, Dragonette was able to derive the general modular transformation formulae by the process of iteration. Her method, however, left certain roots of unity unevaluated and also gave only estimates for the "error" terms arising. In this paper, we shall generalize Watson's method [12; pp. 73-76] and shall thus obtain precise linear transformation formulae with all roots of unity and "error" terms evaluated. This will carry us through Sections II and III. In Section IV, we shall show that all sums of roots of unity that are to be considered are actually generalized Kloosterman sums and therefore subject to the results of Salié [10] and Lehmer [6]. We shall show, in fact, that the exponential sums appearing in the asymptotic expansions of Dragonette [2; Theorems II and IV] are actually special cases of the exponential sum $A_k(n)$ associated with the partition function (cf. [6]). In Section V, we shall apply our results to the problem of improving the error terms in the asymptotic expansion theorems of Dragonette. For example, Dragonette has proved the following [2; p. 478]

THEOREM. *The coefficient of* q^n *in the series* $f(q) = \sum_{n=0}^{\infty} A(n) q^n$ *is*

$$A(n) = \sum_{0<k\leqq n^{\frac{1}{2}}} \frac{\lambda(k) \exp\{\pi(n-1/24)^{\frac{1}{2}}/k\sqrt{6}\}}{k^{\frac{1}{2}}(n-1/24)^{\frac{1}{4}}} + O(n^{\frac{1}{2}} \log n)$$

where

$$\lambda(1) = \tfrac{1}{2}(-1)^{n-1}$$

and

$$\lambda(k) = \tfrac{1}{2} \sum_{\substack{h \bmod 2k \\ \text{g.c.d.}(h,2k)=1}} \exp(-\pi i h n/k)\, \epsilon_{h,k}^{(f)} \qquad (k = 2, 3, \cdots, [n^{\frac{1}{2}}]).$$

We shall improve Dragonette's result to the following

THEOREM 5.1. *The coefficient of* q^n *in the series* $f(q) = \sum_{n=0}^{\infty} A(n)$ *is*

$$A(n) = \sum_{k=1}^{[n^{\frac{1}{2}}]} \frac{\lambda(k) \exp\{\pi(n-1/24)^{\frac{1}{2}}/k\sqrt{6}\}}{k^{\frac{1}{2}}(n-1/24)^{\frac{1}{4}}} + O(n^{\epsilon})$$

where

$$\lambda(k) = \begin{cases} \frac{1}{2}(-1)^{\frac{1}{2}(k+1)}A_{2k}(n) & k \text{ odd} \\ \frac{1}{2}(-1)^{\frac{1}{2}k}A_{2k}(n-\frac{1}{2}k) & k \text{ even,} \end{cases}$$

and ϵ *is positive and arbitrarily small.* $A_k(n)$ *is the exponential sum appearing in the Hardy-Ramanujan formula for* $p(n)$ [9; p. 284].

It seems likely that a closer study of the asymptotic expansion for $A(n)$ would show that the above series would diverge if extended to (since $\lambda(k) = O(k^{\frac{1}{2}+\epsilon})$ and probably $|\lambda(k)| > c \cdot k^{\frac{1}{2}}$ for some $c > 0$ and an infinite number of k). If, however, e^x is replaced by $2\sinh x$ in the above expansion, the problem becomes deeper. I would conjecture that the new series is not absolutely convergent (for probably $|\lambda(p)| > c \cdot p^{\frac{1}{2}}$ for some $c > 0$ and a positive proportion of all primes p), but the extreme accuracy of Dragonette's numerical results [2; p. 494] makes plausible the conjecture that the new series is conditionally convergent and that it converges to $A(n)$.

The author wishes to express his deep indebtedness to Professor Hans Rademacher who suggested this study and gave much advice and encouragement throughout the course of the research.

II. The modular transformation formulae of f and ω. Our first two theorems of this section improve the results in Dragonette's Theorem I [2; p. 477]. The next two theorems improve Dragonette's results for ω in Theorem III [2; p. 499, line (7.7)]. We complete this section by simplifying the expressions for the roots of unity we have obtained.

Throughout the course of this paper we shall endeavor to keep our notation consistent with that of Dragonette. The following well-known theorem together with its notation and conventions will be used over and over.

If $P(q) = \prod_{r=1}^{\infty}(1-q^r)^{-1}$, then

$$\begin{aligned} P(\exp\{2\pi i(h+iz)/k\}) &= \omega_{h,k} z^{\frac{1}{2}} \exp\{\pi(z^{-1}-z)/12k\} \\ &\quad \cdot P(\exp\{2\pi i(h'+iz^{-1})/k\}) \end{aligned} \tag{2.0.1}$$

where $\operatorname{Re} z > 0$ with the principal branch of $z^{\frac{1}{2}}$ taken, $hh' \equiv -1 \pmod{k}$, and $\omega_{h,k}$ a $24k$-th root of unity [9; pp. 290-291]. We remark that $P(q)$ is essentially the modular form $\eta(\tau)$, and (2.0.1) is the transformation formula for $\eta(\tau)$ in thin disguise.

THEOREM 2.1. *Let* g.c.d.$(h,k) = 1$, $hh' + kk' = -1$, h *be even, and* $\operatorname{Re} z > 0$ *with the principal branch of* $z^{\frac{1}{2}}$ *taken. Then*

$$f(\exp\{\pi i(h+iz)/k\}) = \eta_{h,k}{}^{(f)} 2^{\frac{3}{2}} z^{-\frac{1}{2}} \exp(-\pi(4z^{-1}+z/8)/3k)$$
$$\cdot\,\omega(\exp\{2\pi i(h'+iz^{-1})/k\}) + E_1{}^{(f)}(h,k;z)$$

where

$$\eta_{h,k}{}^{(f)} = \exp\{\tfrac{1}{2}\pi i(2h'+k'-3-3h'k'+3h'/k)\}\,\omega_{\frac{1}{2}h,k},$$

and

$$E_1{}^{(f)}(h,k;z) = 2^{\frac{1}{2}}k^{-1}\exp(-\pi z/24k)\sum_{\nu \bmod k}\delta_1{}^{(f)}(\nu;\tfrac{1}{2}h;2h';k)\,z^{\frac{1}{2}}J_1{}^{(f)}(\nu;k;z)$$

with

$$\delta_1{}^{(f)}(\nu;h;h';k) = (-1)^{\nu}\omega_{h,k}\exp\{\pi i h'(-3\nu^2+\nu)/k\}$$

and

$$J_1{}^{(f)}(\nu;k;z) = \int_{-\infty}^{\infty}\frac{\exp(-3\pi z x^2/2k)}{\cosh(\pi i(\nu-\frac{1}{6})/k-\pi z x/2k)}\,dx.$$

Finally if $1 \leqq \nu \leqq k$, *we have*

$$|\,z^{\frac{1}{2}}J_1{}^{(f)}(\nu;k;z)\,| \leqq 3^{-\frac{1}{2}}2^{\frac{3}{2}}N^{\frac{1}{2}}k\,|\,k-2(\nu-\tfrac{1}{6})\,|^{-1}$$

provided $z = kN^{-2}-ik\phi$ *with* $|\,\phi\,| \leqq k^{-1}N^{-1}$ *and* $1 \leqq k \leqq N$.

THEOREM 2.2. *Let* g.c.d.$(h,k)=1$, $hh''+2kk''=-1$, h *be odd, and* $\operatorname{Re} z > 0$ *with the principal branch of* $z^{\frac{1}{2}}$ *taken. Then*

$$f(\exp\{\pi i(h+iz)/k\})$$
$$= \epsilon_{h,k}{}^{(f)}z^{-\frac{1}{2}}\exp(-\pi(z+z^{-1})/24k)\,f(\exp\{\pi i(h''+iz^{-1})/k\}) + E_2{}^{(f)}(h,k;z)$$

where

$$\epsilon_{h,k}{}^{(f)} = (-1)^{k+\frac{1}{2}(hh''-1)}\exp(\pi i h''/2+3\pi i h h''^2 k/2)\,\omega_{h,2k},$$

and

$$E_2{}^{(f)}(h,k;z) = E_1{}^{(f)}(2h,2k;2z)$$
$$= \exp(-\pi z/24k)\,k^{-1}\sum_{\nu \bmod k}\delta_1{}^{(f)}(\nu;h;h'';2k)\,z^{\frac{1}{2}}J_1{}^{(f)}(\nu;2k;2z)$$

with $\delta_1{}^{(f)}(\nu;h;h'';2k)$ *and* $J_1{}^{(f)}(\nu;2k;2z)$ *as in Theorem* 2.1.

THEOREM 2.3. *Let* g.c.d.$(h,k)=1$, $hh'+kk'=-1$, h *be even, and* $\operatorname{Re} z > 0$ *with the principal branch of* $z^{\frac{1}{2}}$ *taken. Then*

$$\omega(\exp\{\pi i(h+iz)/k\}) = \tfrac{1}{2}z^{-\frac{1}{2}}\epsilon_{h,k}{}^{(\omega)}\exp(2\pi z/3k+\pi z^{-1}/12k)$$
$$\cdot\,f(\exp\{2\pi i(h'+iz^{-1})/k\}) + E_1{}^{(\omega)}(h,k;z)$$

where

$$\epsilon_{h,k}^{(\omega)} = (-1)^{\frac{1}{2}(k-1)} \exp(3\pi i h k/4 - 3\pi i h/4k)\, \omega_{h,k}$$

and

$$E_1^{(\omega)}(h,k;z) = k^{-1} \exp(2\pi z/3k) \sum_{\nu \bmod k} \delta_1^{(\omega)}(\epsilon;\nu;h;h';k)\, z^{\frac{1}{2}} J_1^{(\omega)}(\nu;k;z)$$

with

$$\delta_1^{(\omega)}(\epsilon;\nu;h;h';k) = i\epsilon_{h,k}^{(\omega)} \exp(-3\pi i h' \nu^2/k + \pi i h'/k)$$

and

$$J_1^{(\omega)}(\nu;k;z) = \int_{-\infty}^{\infty} \frac{\exp(-3\pi z x^2/k)}{\tanh\{\pi i(\nu - \frac{1}{6})/k - \pi z x/k\}}\, dx.$$

Finally if $-\frac{1}{2}(k+1) \leqq \nu \leqq \frac{1}{2}(k-1)$, *we have*

$$|\, z^{\frac{1}{2}} J_1^{(\omega)}(\nu;k;z) \,| \leqq 2^{-\frac{1}{4}} 3^{-\frac{1}{2}} \exp(\pi/6) N^{\frac{1}{2}} k \,|\, \nu - \tfrac{1}{6} \,|^{-1}$$

provided $z = kN^{-2} - ik\phi$ *with* $|\,\phi\,| \leqq (kN)^{-1}$ *and* $1 \leqq k \leqq N$.

THEOREM 2.4. *Let* g. c. d.$(h,k) = 1$, $hh'' + 2kk'' = -1$, h *be odd, and* $\operatorname{Re} z > 0$ *with the principal branch of* $z^{\frac{1}{2}}$ *taken. Then*

$$\omega(\exp\{\pi i(h + iz)/k\}) = \eta_{h,k}^{(\omega)} z^{-\frac{1}{2}} \exp(2\pi z/3k - 2\pi z^{-1}/3k)$$
$$\cdot\, \omega(\exp\{\pi i(h'' + iz^{-1})/k\}) + E_2^{(\omega)}(h,k;z)$$

where

$$\eta_{h,k}^{(\omega)} = i(-1)^{\frac{1}{2}(h''+1)} \exp\{-3\pi i h'' k''/2 - 3\pi i(h'' - h)/4k\}\, \omega_{h,k}$$

and

$$E_2^{(\omega)}(h,k;z) = k^{-1} \exp(2\pi z/3k) \sum_{\nu \bmod k} (-1)^{\nu} \exp(-3\pi i h''/4k)$$
$$\cdot\, \delta_1^{(\omega)}(\eta;\mu;h;h'';k)\, z^{\frac{1}{2}} J_1^{(\omega)}(\mu;k;z)$$

with $\delta_1^{(\omega)}(\eta;\mu;h;h'';k)$ and $J_1^{(\omega)}(\mu;k;z)$ *as in Theorem* 2.3 *and* $\mu = \nu + \frac{1}{2}$ *written for simplicity of notation.*

We now prove these four theorems. The proof of Theorem 2.1 is given in detail; since the proofs of all four theorems are similar, we omit many of the details in the remaining theorems. Also, it should be noted that Theorem 2.3 may be derived directly from Theorem 2.1; however, such a derivation would be complicated and would not likely shorten our exposition.

Proof of Theorem 2.1. We shall freely utilize the results in [12]. We let

$$
\begin{aligned}
F(q) &= f(q)/P(q) \\
&= f(q)\prod_{r=1}^{\infty}(1-q^r) \\
&= 1+4\sum_{n=1}^{\infty}(-1)^n q^{\frac{1}{2}n(3n+1)}(1+q^n)^{-1} \quad [12;\ \text{p. } 64] \\
&= 1+4\sum_{n=1}^{\infty}(-1)^n q^{3n^2/2}(q^{-\frac{1}{2}n}+q^{\frac{1}{2}n})^{-1}.
\end{aligned} \tag{2.1.1}
$$

If $q=\exp\{\pi i(h+iz)/k\}$, then

$$
\begin{aligned}
F(q) &= 1+2\sum_{n=1}^{\infty}\frac{(-1)^n\exp\{3n^2\pi i(h+iz)/2k\}}{\cosh(n\pi i(h+iz)/2k)} \\
&= \sum_{n=-\infty}^{\infty}\frac{(-1)^n\exp(3\pi ihn^2/2k-3\pi zn^2/2k)}{\cosh(\pi ihn/2k-\pi zn/2k)},\ \text{letting } n=mk+\nu \\
&= \sum_{\nu=0}^{k-1}\sum_{m=-\infty}^{\infty}\frac{(-1)^{mk+\nu}\exp\{3\pi ih(m^2k^2+2mk\nu+\nu^2)/2k-3\pi z(mk+\nu)^2/2k\}}{\cosh\{\pi ihm/2+\pi ih\nu/2k-\pi z(mk+\nu)/2k\}} \\
&= \sum_{\nu=0}^{k-1}(-1)^\nu\exp(3\pi ih\nu^2/2k)\sum_{m=-\infty}^{\infty}\frac{(-1)^m\exp\{-3\pi z(mk+\nu)^2/2k\}}{\cosh\{\pi ih\nu/2k-\pi z(mk+\nu)/2k\}}\ ;
\end{aligned} \tag{2.1.3}
$$

by the Poisson summation formula [1; pp. 7-9] we obtain

$$F(q)=\sum_{\nu=0}^{k-1}(-1)^\nu\exp(3\pi ih\nu^2/2k)\sum_{n=-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\exp\{-3\pi z(kx+\nu)^2/2k+\pi ix+2\pi inx\}}{\cosh\{\pi ih\nu/2k-\pi z(kx+\nu)/2k\}}\,dx;$$

making the change of variable $kx+\nu\rightarrow x$, we obtain .

$$F(q)=k^{-1}\sum_{\nu=0}^{k-1}(-1)^\nu\exp(3\pi ih\nu^2/2k)\sum_{n=-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\exp\{(2n+1)\pi i(x-\nu)/k-3\pi zx^2/2k\}}{\cosh(\pi ih\nu/2k-\pi zx/2k)}\,dx.$$

Notice now that if ν is replaced by $\nu\pm k$ the change introduced is

$$(-1)^{\pm k}\exp\{3\pi ih(\pm 2k+k^2)/2k\}(-1)^{(2n+1)-\frac{1}{2}h}=1.$$

Hence we have

$$F(q)=k^{-1}\sum_{\nu \bmod k}(-1)^\nu\exp(3\pi ih\nu^2/2k)\sum_{n=-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\exp\{(2n+1)\pi i(x-\nu)/k-3\pi zx^2/2k\}}{\cosh(\pi ih\nu/2k-)zx/2k)}\,dx.$$

Now

$$\sum_{n=-\infty}^{\infty}=\sum_{n=0}^{\infty}+\sum_{n=-1}^{-\infty}=\sum_{n=0}^{\infty}+\sum_{n=1}^{\infty}\int_{-\infty}^{\infty}\frac{\exp\{-(2n-1)\pi i(x-\nu)/k-3\pi zx^2/2k\}}{\cosh(\pi ih\nu/2k-)zx/2k)}\,dx;$$

sending $x\rightarrow -x$ and $n\rightarrow n+1$ we derive

$$\sum_{n=-\infty}^{\infty}=\sum_{n=0}^{\infty}+\sum_{n=0}^{\infty}\int_{-\infty}^{\infty}\frac{\exp\{(2n+1)\pi i(x+\nu)/k-3\pi zx^2/2k\}}{\cosh(-\pi ih\nu/2k-\pi zx/2k)}\,dx.$$

Replacing ν by $-\nu$ in the second sum we see that

$$(2.1.4)\quad F(q)=2k^{-1}\sum_{\nu \bmod k}(-1)^{\nu}\exp(3\pi ih\nu^2/2k)$$

$$\cdot\sum_{n=0}^{\infty}\int_{-\infty}^{\infty}\frac{\exp\{(2n+1)\pi i(x-\nu)/k-3\pi zx^2/2k\}}{\cosh(\pi ih\nu/2k-\pi zx/2k)}\,dx.$$

We now define $C_k(x)=\dfrac{\cosh 3x}{\cosh(x/k)}$.

Clearly $C_k(x)$ is an entire function of x if k is an odd integer. Hence if

$$\begin{aligned}\mathfrak{F}_n(x)&=\frac{\exp\{(2n+1)\pi i(x-\nu)/k-3\pi zx^2/2k\}}{\cosh(\pi ih\nu/2k-\pi zx/2k)}\\&=\frac{\exp\{(2n+1)\pi i(x-\nu)/k-3\pi zx^2/2k\}}{\cosh\{3(\pi ih\nu-\pi zx)/2\}}C_k(\{\pi ih\nu-\pi zx\}/2)\\&=\frac{(-1)^{\frac{1}{2}h\nu}\exp\{(2n+1)\pi i(x-\nu)/k-3\pi zx^2/2k\}}{\cosh(3\pi zx/2)}C_k(\{\pi ih\nu-\pi zx\}/2)\end{aligned}$$

then $\mathfrak{F}_n(x)$ has (at most) simple poles at

$$x_m=(2m+1)i/3z,\ m \text{ any integer.}$$

The residue at x_m is

$$\begin{aligned}\lambda_{n,m}=2(3\pi z)^{-1}(-1)^{\frac{1}{2}h\nu-m}\exp\{(2n+1)\pi i(x_m-\nu)/k\\-3\pi zx_m{}^2/2k\}C_k(\{\pi ih\nu-\pi zx_m\}/2).\end{aligned}$$

Now by Cauchy's theorem,

$$\left\{\int_{-\infty}^{\infty}-P\int_{-\infty+(2n+1)i/3z}^{\infty+(2n+1)i/3z}\right\}\mathfrak{F}_n(x)\,dx=2\pi i(\lambda_{n,0}+\lambda_{n,1}+\cdots+\tfrac{1}{2}\lambda_{n,n}).$$

Hence

$$\begin{aligned}F(q)&=2k^{-1}\sum_{\nu \bmod k}(-1)^{\nu}\exp(3\pi ih\nu^2/2k)2\pi i\sum_{n=0}^{\infty}(\lambda_{n,0}+\lambda_{n,1}+\cdots+\tfrac{1}{2}\lambda_{n,n})\\&\quad+2k^{-1}\sum_{\nu \bmod k}(-1)^{\nu}\exp(3\pi ih\nu^2/2k)\sum_{n=0}^{\infty}P\int_{-\infty+x_n}^{\infty+x_n}\mathfrak{F}_n(x)\,dx\\&=\Sigma_1+\Sigma_2.\end{aligned}$$

It will turn out that Σ_1 will yield the main term in the transformation formula involving $\omega(q)$, and Σ_2 will yield the "error" term.

In Σ_1, we have

$$\sum_{n=0}^{\infty}\{\lambda_{n,0}+\lambda_{n,1}+\cdots+\tfrac{1}{2}\lambda_{n,n}\}=\sum_{n=0}^{\infty}\{\tfrac{1}{2}\lambda_{m,m}+\lambda_{m+1,m}+\lambda_{m+2,m}+\cdots\},$$

by rearrangement of the first series,

$$=\sum_{n=0}^{\infty}\lambda_{m,m}\{\tfrac{1}{2}+\sum_{l=1}^{\infty}\exp\{2\pi i(x_m-v)l/k\}$$

$$\text{since } \lambda_{m+l,m}=\exp\{2\pi i(x_m-v)l/k\}\cdot\lambda_{m,m}$$

$$=\tfrac{1}{2}\sum_{n=0}^{\infty}\lambda_{m,m}\frac{1+\exp\{2\pi i(x_m-v)/k\}}{1-\exp\{2\pi i(x_m-v)/k\}}. \tag{2.1.6}$$

We now take a closer look at $\lambda_{m,m}$. Since $\lambda_{m,m}$ is a residue of $\mathfrak{F}_m(x)$, we note that $\lambda_{m,m}\neq 0$ if and only if $\cosh(\pi ihv/2k-\pi zx_m/2k)=0$, i.e. $\cos\{\tfrac{1}{2}\pi(hv/k-(2m+1)/3k\}=0$. Thus we must have $3hv-(2m+1)\equiv 0 \pmod{3k}$.

If $2m+1\not\equiv 0 \pmod 3$, then the above congruence is impossible. Hence $2m+1=6j+3$. Therefore,

$$hv-(2j+1)\equiv 0 \pmod k, \text{ or}$$

$$v\equiv -h'(2j+1) \pmod k$$

$$\text{provided } hh'\equiv -1 \pmod k.$$

Thus when considering $\sum_{v \bmod k}$ we see that the only non-zero term is the term $v\equiv -h'(2j+1) \pmod k$.

We now evaluate $\lambda_{m,m}$ when $m=3j+1$ and $v\equiv -h'(2j+1) \pmod k$.

$$\lambda_{m,m}=\lim_{x\to x_m}(x-x_m)\mathfrak{F}_m(x)$$

$$=\frac{\exp\{(2m+1)\pi i(x_m-v)/k-3\pi zx_m^2/2k\}}{-\tfrac{1}{2}\pi zk^{-1}\sinh(\pi ihv/2k-\pi zx_m/2k)}$$

$$=\frac{\exp\{3(2j+1)\pi i[(2j+1)iz^{-1}+h'(2j+1)]/k+3\pi(2j+1)^2/2kz\}}{-\tfrac{1}{2}\pi zk^{-1}\sinh\{-\pi ihh'(2j+1)/2k-\pi i(2j+1)/2k\}}$$

$$=\frac{-2k\cdot\exp\{3(2j+1)^2\pi i(h'+iz^{-1})/k+3\pi(2j+1)^2/2kz\}}{i\pi z\sin\{\pi(2j+1)kk'/2k\}} \tag{2.1.7}$$

$$=-2k(\pi iz)^{-1}(-1)^{k'j+\frac{1}{2}(k'-1)}\exp\{3(2j+1)^2\pi i(h'+iz^{-1})/k$$
$$+3\pi(2j+1)^2/2kz\}.$$

Hence

$$\Sigma_1=2k^{-1}2\pi i\tfrac{1}{2}\sum_{j=0}^{\infty}(-1)^{-h'(2j+1)}\exp\{3\pi ih(h'(2j+1))^2\}$$
$$\cdot(-2k)(\pi iz)^{-1}(-1)^{k'j+\frac{1}{2}(k'-1)}$$

$$\cdot \exp\{3(2j+1)^2\pi i(h'+iz^{-1})/k + 3\pi(2j+1)^2/2kz\}$$

$$\frac{1+\exp\{2\pi i[(2j+1)i/z + h'(2j+1)]/k\}}{1-\exp\{2\pi i[(2j+1)i/z + h'(2j+1)]/k\}}.$$

$$= 4z^{-1}(-1)^{h'-1+\frac{1}{2}(k'-1)} \sum_{j=0}^{\infty} (-1)^j \exp\{3\pi i h'(2j+1)^2(-1-kk')/2k\}$$

$$(2.1.8) \qquad \cdot \exp\{3(2j+1)^2\pi i(h'+iz^{-1})/k + 3\pi(2j+1)^2/2kz\}$$

$$\frac{1+\exp\{2\pi i[(2j+1)i/z + h'(2j+1)]/k\}}{1-\exp\{2\pi i[(2j+1)i/z + h'(2j+1)]/k\}}.$$

$$= 4z^{-1}(-1)^{h'-1-\frac{1}{2}(k'-1)} \exp\{-3\pi i h'k'/2 + 3\pi i(h'+iz^{-1})/2k\}$$

$$\cdot \sum_{j=0}^{\infty} (-1)^j \exp\{3j(j+1)2\pi i(h'+iz^{-1})/k\}$$

$$\frac{1+\exp\{2\pi i(2j+1)(h'+iz^{-1})/k\}}{1-\exp\{2\pi i(2j+1)(h'+iz^{-1})/k\}}.$$

$$= 4z^{-1}(-1)^{h'-1-\frac{1}{2}(k'-1)} \exp\{-3\pi i h'k'/2 + 3\pi i h'/2k - 3\pi/2kz\}$$

$$\frac{\omega(\exp[2\pi i(h'+iz^{-1})/k])}{P(\exp[4\pi i(h'+iz^{-1})/k])} \quad [12;\ \text{p. } 66].$$

We now evaluate $\mathfrak{S}_2$. We recall that

$$\mathfrak{S}_2 = 2k^{-1} \sum_{\nu \bmod k} (-1)^{\nu} \exp(3\pi i h\nu/2k) \sum_{n=0}^{\infty} P\int_{-\infty+x_n}^{\infty+x_n} \mathfrak{F}_n(x)\,dx,$$

where $x = (2n+1)i/3z$.

Now

$$2\sum_{n=0}^{\infty} = \sum_{n=0}^{\infty} + \sum_{n=0}^{-\infty} P\int_{-\infty-(2n+1)i/3z}^{\infty+(2n+1)i/3z} \frac{\exp\{-(2n-1)\pi i(x-\nu)/k - 3\pi z x^2/2k\}}{\cosh(\pi i h\nu/2k - \pi z x/2k)}\,dx$$

$$= \sum_{n=0}^{\infty} + \sum_{n=1}^{\infty-} P\int_{-\infty+(-2n+1)i/3z}^{\infty+(-2n+1)i/3z} \frac{\exp\{(2n+1)\pi i(x-\nu)/k - 3\pi z x^2/2k\}}{\cosh(\pi i h\nu/2k - \pi z x/2k)}\,dx$$

upon replacing ν by $-\nu$ and sending $x \to -x$; thus

$$\mathfrak{S}_2 = k^{-1} \sum_{\nu \bmod k} (-1)^{\nu} \exp(3\pi i h\nu^2/2k) \sum_{n=-\infty}^{\infty} P\int_{-\infty+x_n}^{\infty+x_n} \mathfrak{F}_n(x)\,dx.$$

Now

$$P\int_{-\infty+x_n}^{\infty+x_n} \mathfrak{F}_n(x)\,dx = P\int_{-\infty}^{\infty} \mathfrak{F}_n(x+x_n)\,dx,$$

and

$$\exp\{(2n+1)\pi i(x+x_n-\nu)/k - 3\pi z(x+x_n)^2/2k\}$$

$$\begin{aligned} &= \exp\{(2n+1)\pi ix/k - (2n+1)^2\pi/3kz - (2n+1)\pi i\nu/k \\ &\qquad - 3\pi zx^2/2k - \pi x(2n+1)i/k\} \cdot \exp\{\pi(2n+1)^2/6kz\} \\ &= \exp\{-(2n+1)\pi i\nu/k - (2n+1)^2\pi/6kz - 3\pi zx^2/2k\} \end{aligned}$$

We return to Σ_2 letting $n = 3p + \delta$,

$$\begin{aligned} \Sigma_2 = k^{-1} \sum_{\nu \bmod k} (-1)^\nu \exp(3\pi ih\nu^2/2k) \sum_{\delta=-1}^{1} \sum_{p=-\infty}^{\infty} \exp\{&-(2n+1)\pi i\nu/k \\ &-(2n+1)^2\pi/6kz\} \end{aligned} \tag{2.1.9}$$

$$\cdot P\int_{-\infty}^{\infty} \frac{\exp(-3\pi zx^2/2k)}{\cosh\{\pi ih\nu/2k - \pi zx/2k - \pi ip/k - \pi i(2\delta+1)/6k\}}\,dx$$

If we replace $\frac{1}{2}h - p$ by ν, i.e. $\nu \to -2h'\nu - 2ph'$, then

$$\begin{aligned} &\cosh\{\pi i(\tfrac{1}{2}h\nu - p)/k - \pi zx/2k - \pi i(2\delta+1)/6k\} \\ &\quad \to \cosh\{-\pi ip/k - \pi i(1+kk')(\nu+p)/k - \pi zx/2k - \pi i(2\delta+1)/6k\} \\ &\qquad = (-1)^{\nu+p}\cosh\{\pi i\nu/k - \pi zx/2k - \pi i(2\delta+1)/6k\}, \end{aligned}$$

and

$$\begin{aligned} (-1)^\nu \exp\{3\pi ih\nu^2/2k - (2n+1)\pi i\nu/k\} &\to \exp\{3\pi ih4h'^2(\nu+p)^2 \\ &\qquad + 2(6p+2\delta+1)\pi ih'(\nu+p)/k\} \\ = \exp\{-6\pi ih'(\nu^2+2\nu p+p^2)/k &+ 12p\pi ih'(\nu+p)/k \\ &+ 2(2\delta+1)\pi ih'p/k + 2(2\delta+1)\pi ih'\nu/k\} \\ = \exp\{6\pi ih'p^2/k + (2\delta+1)&\pi i2h'p/k \\ &- 6\pi ih'\nu^2/k + (2\delta+1)\pi i2h'\nu/k\}. \end{aligned}$$

Thus in Σ_2 we now have $\sum_{p=-\infty}^{\infty}$ free of x and ν. We define

$$\begin{aligned} U_\delta(z) &= \sum_{p=-\infty}^{\infty} (-1)^p \exp\{6\pi ih'p^2/k + (2\delta+1)\pi ih'p/k \\ &\qquad\qquad - (6p+2\delta+1)^2\pi/6kz\} \\ &= \exp\{-(2\delta+1)^2\pi/6kz\} \sum_{p=-\infty}^{\infty} (-1)^p \exp\{2\pi i(h'+iz^{-1}) \\ &\qquad\qquad \cdot (3p^2+(2\delta+1)p)/k\}. \end{aligned} \tag{2.1.10}$$

Letting $\tau' = (h'+iz^{-1})/k$, we have for $\delta = 1$

$$U_1(z) = \exp(-3\pi/2kz) \sum_{p=-\infty}^{\infty} (-1)^p \exp\{2\pi i\tau' 3p(p+1)\}.$$

In this series the p-th term and the $(-p-1)$-st term cancel each other, hence $U_1(z)=0$.

On the other hand,

$$\begin{aligned} U_0(z) &= \exp(-\pi/6kz)\sum_{p=-\infty}^{\infty}(-1)^p\exp\{2\pi i r' p(3p+1)\} \\ &= \exp(-\pi/6kz)\sum_{p=-\infty}^{\infty}(-1)^p\exp\{2\pi i r' p(3p-1)\} \\ &= U_{-1}(z) \\ &= \frac{\exp(-\pi/6kz)}{P(\exp[4\pi i(h'+iz^{-1})/k])} \qquad [4;\text{ p. } 284]. \end{aligned} \tag{2.1.11}$$

Hence

$$\Sigma_2 = \frac{k^{-1}\exp(-\pi/6kz)}{P[4\pi i(h'+iz^{-1})/k]}\sum_{\delta=-1}^{0}\sum_{\nu \bmod k}(-1)^\nu\exp\{-6\pi i h'\nu^2/k + (2\delta+1)\pi i 2h'\nu/k\} \cdot \int_{-\infty}^{\infty}\frac{\exp(-3\pi z x^2/2k)}{\cosh\{\pi i\nu/k - \pi z x/2k - \pi i(2\delta+1)/6k\}}\,dx.$$

In the case $\delta=-1$, if we replace ν by $-\nu$ and send $x\to -x$ we get the case $\delta=0$, hence

$$\Sigma_2 = \frac{k^{-1}2\cdot\exp(-\pi/6kz)}{P[4\pi i(h'+iz^{-1})/k]}\sum_{\nu \bmod k}(-1)^\nu\exp\{-6\pi i h'\nu^2/k + 2\pi i h'\nu/k\} \cdot \int_{-\infty}^{\infty}\frac{\exp(-3\pi z x^2/2k)}{\cosh\{\pi i(\nu-\frac{1}{6})/k - \pi z x/2k\}}\,dx. \tag{2.1.12}$$

Since h is even, we may write (2.0.1) as

$$P(\exp\{2\pi i(\tfrac{1}{2}h + i\tfrac{1}{2}z)/k\}) = \omega_{\frac{1}{2}h,k}2^{-\frac{1}{2}}z^{\frac{1}{2}}\exp(\pi/6kz - \pi z/24k)P(\exp\{4\pi i(h'+iz^{-1})/k\});$$

thus since we have established

$$\frac{f(\exp\{\pi i(h+iz)/k\})}{P(\exp\{\pi i(h+iz)/k\})} = \Sigma_1 + \Sigma_2,$$

we have

$$\begin{aligned} &f(\exp[\pi i(h+iz)/k]) \\ &\quad = (-1)^{h'-1+\frac{1}{2}(k'-1)}\exp\{-3\pi i h'k'/2 + 3\pi i h'/2k\}\omega_{\frac{1}{2}h,k}2^{\frac{3}{2}}z^{-\frac{1}{2}} \\ &\quad\quad \cdot\exp(-4\pi/3kz - \pi z/24k)\omega(\exp[2\pi i(h'+iz^{-1})/k]) \end{aligned}$$

$$(2.1.13) \qquad + 2^{\frac12} z^{\frac12} k^{-1} \exp(-\pi z/24k)\,\omega_{\frac12 h,k} \sum_{\nu \bmod k} (-1)^\nu \exp(-6\pi i h'\nu^2/k - 2\pi i h'\nu/k) \cdot \int_{-\infty}^{\infty} \frac{\exp(-3\pi z x^2/2k)}{\cosh\{\pi i(\nu - \frac16)/k - \pi z x/2k\}}\, dx.$$

This is the required transformation formula; thus when h is even, the general linear transformation sends f into ω plus a correction $E_1^{(f)}(h, k; z)$.

We proceed to the final part of the theorem in which we estimate the integral involved in $E_1^{(f)}(h, k; z)$.

$$z^{\frac12} J_1^{(f)}(\nu; k; z) = z^{\frac12} \int_{-\infty}^{\infty} \frac{\exp(-3\pi z x^2/2k)}{\cosh\{\pi i(\nu - \frac16)/k - \pi z k/2k\}}\, dx$$
$$= 2k\pi^{-1} z^{-\frac12} \int_{-\infty}^{\infty} \frac{\exp(-6kt^2/\pi z)}{\cosh\{\pi i(\nu - \frac16)/k - t\}}\, dt$$

by Cauchy's theorem. Now with $z = kN^{-2} - ik\phi$ and $1 \leqq \nu \leqq k$, we have

$$| z^{\frac12} J_1^{(f)}(\nu; k; z) | \leqq 2k\pi^{-1} | z |^{-\frac12} \int_{-\infty}^{\infty} \frac{\exp\{-6k(\operatorname{Re} z^{-1})t^2/\pi\}}{| \cosh\{\pi i(\nu - \frac16)/k - t\} |}\, dt$$
$$\leqq 2k\pi^{-1} | z |^{-\frac12} | \cos\{\pi(\nu - \tfrac16)/k\} |^{-1} \int_{-\infty}^{\infty} \exp\{-6k(\operatorname{Re} z^{-1})t^2/\pi\}\, dt$$
$$\leqq 2k\pi^{-1} | z |^{-\frac12} | 1 - 2k^{-1}(\nu - \tfrac16) |^{-1} \pi^{\frac12} \{12k(\operatorname{Re} z^{-1})\}^{-\frac12} \int_{-\infty}^{\infty} \exp(-u^2/2)\, du$$
$$\leqq k^{\frac32} (3\pi)^{-\frac12} | k - 2(\nu - \tfrac16) |^{-1} | z |^{-\frac12} (\operatorname{Re} z^{-1})^{-\frac12} (2\pi)^{\frac12}$$
$$= k^{\frac32} (\tfrac23)^{\frac12} | k - 2(\nu - \tfrac16) |^{-1} | z |^{-\frac12} \{(\operatorname{Re} z)/| z |^2\}^{-\frac12}.$$

Now $\operatorname{Re} z = kN^{-2}$, and

$$| z | = (k^2 N^{-4} + k^2\phi^2)^{\frac12}$$
$$\leqq (k^2 N^{-4} + k^2 k^{-2} N^{-2})^{\frac12}$$
$$= \{(k^2 + N^2) N^{-4}\}^{\frac12}$$
$$\leqq 2^{\frac12} N^{-1}.$$

Hence

$$| z |^{\frac12} (\operatorname{Re} z)^{-\frac12} \leqq 2^{\frac14} N^{-\frac12} N k^{-\frac12} = 2^{\frac14} N^{\frac12} k^{-\frac12}.$$

Thus

$$| z^{\frac12} J_1^{(f)}(\nu; k; z) | \leqq 2^{\frac34} N^{\frac12} k 3^{-\frac12} | k - 2(\nu - \tfrac16) |^{-1}$$

This concludes the proof of Theorem 2.1.

Proof of Theorem 2.2. As before

$$F(q) = f(q)/P(q) \tag{2.2.1}$$

$$= 1 + 4\sum_{n=1}^{\infty} \frac{(-1)^n q^{\frac{1}{2}n(3n+1)}}{1+q^n} \tag{2.2.2}$$

$$= \sum (-1)^\nu \exp(3\pi i h\nu^2/2) \sum_{\nu=0}^{2k-1} \frac{(-1)^m \exp\{-3\pi z(2mk+\nu)^2/2k\}}{\cosh\{\pi i h\nu/2k - \pi z(2mk+\nu)/2k\}} \tag{2.2.3}$$

by the Poisson summation formula [1; pp. 7-9] we obtain

$$\begin{aligned} F(q) &= k^{-1} \sum_{\nu \bmod 2k} (-1)^\nu \\ &\cdot \exp(3\pi i h\nu^2/2k) \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \frac{\exp\{(2n+1)\pi i(x-\nu)/2k - 3\pi z x^2/2k\}}{\cosh(\pi i h\nu/2k - \pi z x/2k)} dx. \end{aligned} \tag{2.2.4}$$

We now define

$$S_{2k}(x) = \frac{\sinh 3x}{\cosh(x/2k)}.$$

Clearly $S_{2k}(x)$ is an entire function of x since k is an integer. Hence if

$$\begin{aligned} \mathfrak{F}_n(x) &= \frac{\exp\{(2n+1)\pi i(x-\nu)/2k - 3\pi z x^2/2k\}}{\cosh(\pi i h\nu/2k - \pi z x/2k)} \\ &= \frac{-(-1)^\nu \exp\{(2n+1)\pi i(x-\nu)/2k - 3\pi z x^2/2k\} S_{2k}(\pi i h\nu - \pi z x)}{\sinh 3\pi z x}, \end{aligned}$$

then $\mathfrak{F}_n(x)$ has (at most) simple poles at

$$x_m = 2mi/6z = mi/3z.$$

The residue at x_m is

$$\begin{aligned} \lambda_{n,m} = -(-1)^{\nu+m}(3\pi z)^{-1}\exp\{(2n+1)\pi i(x_m-\nu)/2k \\ - 3\pi z x_m{}^2/2k\} S_{2k}(\pi i h\nu - \pi z x_m). \end{aligned}$$

Now by Cauchy's theorem,

$$\left\{\int_{-\infty}^{\infty} - \int_{-\infty+(2n+1)i/6z}^{\infty+(2n+1)i/6z}\right\} \mathfrak{F}_n(x)\,dx = 2\pi i\{\tfrac{1}{2}\lambda_{n,0} + \lambda_{n,1} + \cdots + \lambda_{n,n}\}.$$

Hence

$$\begin{aligned} F(q) &= k^{-1} \sum_{\nu \bmod 2k} (-1)^\nu \exp(3\pi i h\nu^2/2k) \sum_{n=0}^{\infty} 2\pi i\{\tfrac{1}{2}\lambda_{n,0} + \lambda_{n,1} + \cdots + \lambda_{n,n}\} \\ &\quad + k^{-1} \sum_{\nu \bmod 2k} (-1)^\nu \exp(3\pi i h\nu^2/2k) \sum_{n=0}^{\infty} \int_{-\infty+(2n+1)i/6z}^{\infty+(2n+1)i/6z} \mathfrak{F}_n(x)\,dx \\ &= \Sigma_1 + \Sigma_2. \end{aligned} \tag{2.2.5}$$

In this instance, Σ_1 will yield the main term in the transformation formula involving $f(q)$, and Σ_2 will again yield the "error" term.

In Σ_1, we have

$$\sum_{n=0}^{\infty}\{\tfrac{1}{2}\lambda_{n,0}+\lambda_{n,1}+\cdots+\lambda_{n,n}\}$$

$$= \frac{1}{2}\frac{\lambda_{0,0}}{1-\exp\{\pi i(x_m-\nu)/k\}}+\sum_{m=1}^{\infty}\frac{\lambda_{m,m}}{1-\exp\{\pi i(x_m-\nu)/k\}}. \tag{2.2.6}$$

We thus obtain

$$\Sigma_1=\pi i k^{-1}\sum_{\nu \bmod 2k}(-1)^{\nu}\exp(3\pi i h\nu^2/2k)\sum_{m=-\infty}^{\infty}\frac{\lambda_{m,m}}{1-\exp\{\pi i(x_m-\nu)/k\}}. \tag{2.2.6a}$$

We now take a closer look at $\lambda_{m,m}$. Since $\lambda_{m,m}$ was a residue of $\mathfrak{F}_m(x)$, we note that $\lambda_{m,m}\neq 0$ if and only if $\cosh(\pi i h\nu/2k-\pi z x_m/2k)=0$, i.e. $\cos\{\pi(h\nu/k-m/3k)/2\}=0$. Thus we must have $3h\nu-m\equiv 3k \pmod{6k}$. If $m\not\equiv 0 \pmod 3$, then this last congruence is impossible. Hence $m=3j$.

$$h\nu-j\equiv k \pmod{2k}$$

$$\text{or} \qquad \nu\equiv -h''(k+j) \pmod{2k}$$

$$\text{provided } hh''\equiv -1 \pmod{2k}.$$

Thus when considering $\sum\limits_{\nu \bmod 2k}$, we see that the only non-zero term is the term $\nu\equiv -h''(k+j) \pmod{2k}$.

We now evaluate $\lambda_{m,m}$ with $m=3j$ and $\nu\equiv -h''(k+j) \pmod{2k}$.

$$\lambda_{m,m}=\lim_{x\to x_m}(x-x_m)\mathfrak{F}_m(x)$$

$$=2k(\pi i z)^{-1}(-1)^{jk''+\frac{1}{2}(hh''-1)+j}\exp\{\pi i h''/2 + (6j+1)\pi i j(h''+iz^{-1})/2k+3\pi j^2/2kz\} \tag{2.2.7}$$

where $hh''+2kk''=-1$. Hence

$$\Sigma_1=z^{-1}(-1)^{k+\frac{1}{2}(hh''-1)}\exp(\pi i h''/2 + 3\pi i h h''^2 k/2)\frac{f(\exp\{\pi i(h''+iz^{-1})/k\})}{P(\exp\{\pi i(h''+iz^{-1})/k\})} \text{ by [12; p. 64].} \tag{2.2.8}$$

We now evaluate Σ_2.

$$\Sigma_2=k^{-1}\sum_{\nu \bmod 2k}(-1)^{\nu}\exp(3\pi i h\nu^2/2k)\sum_{n=0}^{\infty}\int_{-\infty+(2n+1)i/6z}^{\infty+(2n+1)i/6z}\mathfrak{F}_n(x)\,dx$$

$$=\tfrac{1}{2}k^{-1}\sum_{\nu \bmod 2k}(-1)^{\nu}\exp(3\pi i h\nu^2/2k)\sum_{n=-\infty}^{\infty}\int_{-\infty+(2n+1)i/6z}^{\infty+(2n+1)i/6z}\mathfrak{F}_n(x)\,dx,$$

as in Theorem 2.1.

Now

$$\int_{-\infty+(2n+1)i/6z}^{\infty+(2n+1)i/6z} \mathcal{F}_n(x)\,dx = \int_{\infty}^{-\infty} \mathcal{F}_n(x(2n+1)i/6z)\,dx,$$

and in Σ_2 letting $n = 3p + \delta$, we have

$$\Sigma_2 = \tfrac{1}{2}k^{-1} \sum_{h \bmod 2k} (-1)^\nu$$

$$(2.2.9) \qquad \cdot \exp(3\pi i h\nu^2/2k) \sum_{\delta=-1}^{1} \sum_{p=-\infty}^{\infty} \exp\{-(2n+1)\pi i\nu/2k\}$$

$$\cdot \exp\{-(2n+1)^2\pi/24kz\}$$

$$\int_{-\infty}^{\infty} \frac{\exp(-3\pi z x^2/2k)}{\cosh\{\pi i h\nu/2k - \pi z x/2k - \pi i(6p+2\delta+1)/12k\}} dx$$

If we replace $h\nu - p$ by ν i.e. $\nu \rightarrow -h''\nu - h''p$, then in Σ_2 we now have $\sum_{p=-\infty}^{\infty}$ free of x and ν. We define

$$(2.2.10) \qquad U_\delta(z) = \sum_{p=-\infty}^{\infty} (-1)^p \exp\{3\pi i h'' p^2/2k + (2\delta+1)\pi i h'' p/2k - (6p+2\delta+1)^2\pi/24kz\}$$

$$(2.2.11) \qquad = \begin{cases} 0 \text{ if } \delta = 1 \\ \exp(-\pi/24kz) \dfrac{1}{P(\exp\{\pi i(h'' + iz^{-1})/k\})} \end{cases}$$

as in Theorem 2.1. Hence

$$\Sigma_2 = \frac{k^{-1}\exp(-\pi/24kz)}{P(\exp\{\pi i(h'' + iz^{-1})/k\})} \sum_{\nu \bmod 2k} (-1)^\nu$$

$$(2.2.12) \qquad \cdot \exp(-3\pi i h''\nu^2/2k + \pi i h''\nu/2k)$$

$$\cdot \int_{-\infty}^{\infty} \frac{\exp(-3\pi z x/2k)}{\cosh(\pi i h\nu/2k - \pi z x/2k - \pi i/12k)} dx.$$

Since h is odd, we may write (2.0.1) as

$$P(\exp\{2\pi i(h + iz)/2k\})$$

$$= \omega_{h,2k} z^{\frac{1}{2}} \exp(\pi/24kz - \pi z/24k) P(\exp\{2\pi i(h'' + iz^{-1})/2k\});$$

thus since we have established

$$\frac{f(\exp\{\pi i(h+iz)/k\})}{P(\exp\{\pi i(h+iz)/k\})} = \Sigma_1 + \Sigma_2,$$

we have

$$
\begin{aligned}
&f(\exp\{\pi i(h+iz)/k\}) \\
&\quad = (-1)^{k+\frac{1}{2}(hh''-1)}\exp(\pi ih''/2 + 3\pi ihh''^2k/2)\omega_{h,2k}z^{-\frac{1}{2}} \\
&\qquad \cdot \exp(\pi/24kz - \pi z/24k)f(\exp\{\pi i(h'' + iz^{-1})/k\}) \\
&\quad + k^{-1}\omega_{h,2k}z^{\frac{1}{2}}\exp(-\pi z/24k) \sum_{\nu \bmod 2k} (-1)^{\nu} \\
&\qquad \cdot \exp(-3\pi ih''\nu^2/2k + \pi ih''\nu/2k) \\
&\qquad \cdot \int_{-\infty}^{\infty} \frac{\exp(-3\pi zx^2/2k)}{\cosh(\pi i\nu/2k - \pi zx/2k - \pi i/12k)}\,dx.
\end{aligned} \tag{2.2.13}
$$

Thus when h is odd, the general linear transformation sends f into f plus a correction $E_2^{(f)}(h,k;z)$; this completes Theorem 2.2.

Proof of Theorem 2.3. We let

$$\Omega(q) = \omega(q)/P(q^2) \tag{2.3.1}$$

$$= q^{-\frac{3}{4}}\sum_{n=0}^{\infty}(-1)^n q^{3(n+\frac{1}{2})^2}\frac{1+q^{2n+1}}{1-q^{2n+1}}, \text{ [12; p. 66]}, \tag{2.3.2}$$

and if $q = \exp\{\pi i(h+iz)/k\}$

$$
\begin{aligned}
&q^{\frac{3}{4}}\Omega(q) \\
&= -\tfrac{1}{2}\sum_{\nu=0}^{k-1}\sum_{m=-\infty}^{\infty}\frac{(-1)^{\nu+m}\exp\{3(\nu+\frac{1}{2})^2\pi ih/k - 3\pi z(nk+\nu+\frac{1}{2})^2/k\}}{\tanh\{(\nu+\frac{1}{2})\pi ih/k - (mk+\nu+\frac{1}{2})\pi z/k\}};
\end{aligned} \tag{2.3.3}
$$

by the Poisson summation formula [1; pp. 7-9] we obtain

$$
\begin{aligned}
q^{\frac{3}{4}}\Omega(q) = -k^{-1}\sum_{\nu \bmod k}(-1)^{\nu}\exp(3\pi ih\mu^2/k)\sum_{n=0}^{\infty} \\
\cdot\int_{-\infty}^{\infty}\frac{\exp\{(2n+1)\pi i(x-\mu)/k - 3\pi zx^2/k\}}{\tanh(\pi ih\mu/k - \pi zx/k)}dx
\end{aligned} \tag{2.3.4}
$$

as in the previous theorems with $\mu = \nu + \frac{1}{2}$. Hence if $\mathcal{S}_k(x) = \coth(x/k) \cdot \sinh 3x$, then we have $\mathcal{S}_k(x)$ is an entire function of x since k is an integer, and

$$
\begin{aligned}
\mathcal{F}_n(x) &= \frac{\exp\{(2n+1)\pi i(x-\mu)/k - 3\pi zx^2/k\}}{\tanh(\pi ih\mu/k - \pi zx/k)} \\
&= \frac{-\exp\{(2n+1)\pi i(x-\mu)/k - 3\pi zx^2/k\}\mathcal{S}_k(\pi ih\mu - \pi zx)}{(-1)^{\frac{1}{2}h}\sinh 3\pi zx}.
\end{aligned}
$$

Thus $\mathcal{F}_n(x)$ has (at most) simple poles at

$$x_m = mi/3z = 2mi/6z.$$

The residue at x_m is

$$\lambda_{n,m} = -(-1)^{m+\frac{3}{2}h}(3\pi z)^{-1}\exp\{(2n+1)\pi i(x_m-\mu)/k$$
$$-3\pi z x_m^2/k\mathcal{S}_k(\pi ih\mu - \pi z x_m),$$

and

$$\lambda_{m+l,m} = \exp\{2\pi il(x_m-\mu)/k\}\lambda_{m,m}.$$

Now by Cauchy's theorem,

$$\left\{\int_{-\infty}^{\infty} - \int_{-\infty+(2n+1)i/6z}^{\infty+(2n+1)i/6z}\right\}\mathcal{F}_n(x)dx = 2\pi i\{\tfrac{1}{2}\lambda_{n,0}+\lambda_{n,1}+\cdots+\lambda_{n,n}\}.$$

Hence

$$q^{\frac{3}{4}}\Omega(q) = -2\pi ik^{-1}\sum_{\nu \bmod k}(-1)^{\nu}$$
$$\cdot\exp(3\pi ih\mu^2/k)\sum_{n=0}^{\infty}\{\tfrac{1}{2}\lambda_{n,0}+\lambda_{n,1}+\cdots+\lambda_{n,n}\}$$
$$(2.3.5) \qquad -k^{-1}\sum_{\nu \bmod k}(-1)^{\nu}\exp(3\pi ih\mu^2/k)\sum_{n=0}^{\infty}\int_{-\infty+(2n+1)i/6z}^{\infty+(2n+1)i/6z}\mathcal{F}_n(x)\,dx.$$
$$=\Sigma_1+\Sigma_2.$$

Again Σ_1 yields the main term of the transformation this time involving $f(q)$, and Σ_2 as always yields the "error" term.

In Σ_1 we have

$$\sum_{n=0}^{\infty}\{\lambda_{n,0}+\cdots+\lambda_{n,n}\}$$
$$(2.3.6) \qquad = \frac{1}{2}\frac{\lambda_{0,0}}{1-\exp\{2\pi i(x_m-\mu)/k\}} + \sum_{m=1}^{\infty}\frac{\lambda_{m,m}}{1-\exp\{2\pi i(x_m-\mu)/k\}};$$

so

$$\Sigma_1 = -\pi ik^{-1}\sum_{\nu \bmod k}(-1)^{\nu}$$
$$(2.3.6a) \qquad \cdot\exp(3\pi ih\mu^2/k)\sum_{m=-\infty}^{\infty}\frac{\lambda_{m,m}}{1-\exp\{2\pi i(x_m-\mu)/k\}}.$$

We now take a closer look at $\lambda_{m,m}$. Since $\lambda_{m,m}$ was a residue of $\mathcal{F}_m(x)$ we note that $\lambda_{m,m}\neq 0$ if and only if $\tanh(\pi ih\mu/k - \pi zx/k) = 0$, i.e. $\sin\{h\mu/k - m/3k\} = 0$. Thus we must have

$$3h(2\nu+1) - 2m \equiv 0 \pmod{6k}$$

since μ is the abbreviation for $\nu+\frac{1}{2}$. If $m \not\equiv 0 \pmod 3$, then the above congruence is impossible. Hence $m = 3j$.

$$\tfrac{1}{2}h(2\nu+1)-j\equiv 0 \pmod{k}$$

or

$$2\nu+1\equiv -2h'j \pmod{k}$$

or

$$\nu\equiv\tfrac{1}{2}(k+1)(-2h'j-1) \pmod{k}$$
$$\equiv -h'j-\tfrac{1}{2}(k+1) \pmod{k};$$

provided $hh'\equiv -1 \pmod{k}$.

Thus when considering $\sum\limits_{\nu \bmod k}$ we see that the only non-zero term is the term $\nu\equiv -h'j-\tfrac{1}{2}(k+1) \pmod{k}$.

We now evaluate $\lambda_{m,m}$ when $m=3j$ and $\nu\equiv -h'j-\tfrac{1}{2}(k+1) \pmod{k}$.

$$\lambda_{m,m}=\lim_{x\to x_m}(x-x_m)\mathfrak{F}_m(x)$$

(2.3.7)

$$=-ik(\pi z)^{-1}\exp\{\pi ij(h'+iz^{-1})/\mathrm{k}-3\pi j^2/kz+6\pi ih'j^2/k\}.$$

Hence

(2.3.8) $$\Sigma_1=(2z)^{-1}(-1)^{\frac{1}{2}(k-1)}\exp(3\pi ihk/4)F(\exp\{2\pi i(h'+iz^{-1})/k\}),$$

where $F(q)$ is defined in the first line of the proof of Theorem 2.1.

We now evaluate Σ_2. With $n=3p+\delta$ and $\mu=\nu+\tfrac{1}{2}$, we have

(2.3.9)

$$\Sigma_2=-(2k)^{-1}\sum(-1)^{\nu}\exp(3\pi ih\mu^2/k)\sum_{\delta=-1}^{1}\sum_{p=-\infty}^{\infty}$$
$$\cdot\int_{-\infty}^{\infty}\frac{\exp\{-(2n+1)\pi i\mu/k-(2n+1)^2\pi/12kz-3\pi zx^2/k\}}{\tanh\{\pi ih\mu/k-\pi zx/k-\pi i(6p+2\delta+1)/6k\}}dx.$$

If we replace $h\mu-p$ by ν, i.e. $\nu\to -h'(\nu+p)-\tfrac{1}{2}(k+1)$, then we have $\sum\limits_{p=-\infty}^{\infty}$ free of x and ν. We define

(2.3.10)

$$U_\delta(z)=\sum_{p=-\infty}^{\infty}(-1)^p\exp\{\pi ih'(3p^2+(2\delta+1)p)/k$$
$$-(6p+2\delta+1)^2\pi/12kz\}.$$

As before

$$U_1(z)=0$$

(2.3.11) $$U_{-1}(z)=U_0(z)=\frac{\exp(-\pi/12kz)}{P(\exp\{2\pi i(h'+iz^{-1})/k\})}.$$

Thus

$$\Sigma_2=\frac{ik^{-1}(-1)^{\frac{1}{2}(k-1)}\exp(3\pi ihk/4-\pi/12kz)}{P(\exp\{2\pi i(h'+iz^{-1})/k\})}\sum_{\nu \bmod k}exp(-3\pi ih'\nu^2/k)$$

(2.3.12)
$$\cdot \exp(\pi i h'\nu/k) \int_{-\infty}^{\infty} \frac{\exp(-3\pi z x^2/k)}{\tanh(\pi i\nu/k - \pi z x/k - \pi i/6k)}\,dx.$$

We have established

$$\exp(3\pi i(h+iz)/4k)\,\frac{\omega(\exp\{\pi i(b+iz)/k\})}{P(\exp\{2\pi i(h+iz)/k\})} = \Sigma_1 + \Sigma_2,$$

hence by (2.0.1)

$$\begin{aligned}
&\omega(\exp\{\pi i(h+iz)/k\})\\
(2.3.13)\quad &= \tfrac{1}{2} z^{-\frac12}(-1)^{\frac12(k-1)}\exp(3\pi i h k/4 - 3\pi i h/4k)\,\omega_{h,k}\exp(\pi/12kz + 2\pi z/3k)\\
&\qquad \cdot f(\exp\{2\pi i(h'+iz^{-1})/k\})\\
&\quad + i(-1)^{\frac12(k-1)}\exp(3\pi i h k/4 - 3\pi i h/4k + 2\pi z/3k)\,k^{-1}z^{\frac12}\omega_{h,k}\\
&\quad \cdot \sum_{\nu \bmod k} \exp(-3\pi i h'\nu^2/k + \pi i h'\nu/k) \int_{-\infty}^{\infty} \frac{\exp(-3\pi z x^2/k)}{\tanh(\pi i(\nu - \frac16)/k - \pi z x/k)}\,dx.
\end{aligned}$$

This is the required transformation formula.

We proced to the final part of the theorem in which we estimate the integral appearing in $E_1^{(\omega)}(h,k;z)$.

$$\begin{aligned}
z^{\frac12}J_1^{(\omega)}(\nu;k;z) &= z^{\frac12}\int_{-\infty}^{\infty} \frac{\exp(-3\pi z x^2/k)}{\tanh(\pi i(\nu-\frac16)/k - \pi z x/k)}\,dx\\
&= k(\pi z^{\frac12})^{-1}\int_{-\infty}^{\infty}\frac{\exp(-3kt^2/\pi z)}{\tanh(\pi i(\nu-\frac16)/k - t)}\,dt
\end{aligned}$$

by Cauchy's theorem. Now with $z = kN^{-2} - ik\phi$ and $-\frac12(k+1) \leqq \nu \leqq \frac12(k-1)$, we have

$$\begin{aligned}
|z^{\frac12}J_1^{(\omega)}(\nu;k;z)| &\leqq k\pi^{-1}|z|^{-\frac12}\\
&\int_{-\infty}^{\infty} \frac{\exp(-3k(\operatorname{Re} z^{-1})t^2/\pi)\,|\cosh\{\pi i(\nu-\frac16)/k - t\}|}{|\sinh\{\pi i(\nu-\frac16)/k - t\}|}\,dt.
\end{aligned}$$

Now
$$|\cosh(x+iy)| = (\cos^2 y\cosh^2 x + \sin^2 y \sinh^2 x)^{\frac12} \leqq \sqrt{2}\cosh x,$$
and
$$|\sinh(x+iy)| = (\sinh^2 x\cos^2 y + \cosh^2 x \sin^2 y)^{\frac12} \geqq |\sin y|.$$
Thus
$$\begin{aligned}
|z^{\frac12}J_1^{(\omega)}(\nu;k;z)| &\leqq k\sqrt{2}\,\pi^{-1}|z|^{-\frac12}|\sin\{\pi(\nu-\tfrac16)/k\}|^{-1}\int_{-\infty}^{\infty}\\
&\qquad \cdot \exp(-3k(\operatorname{Re} z^{-1})t^2/\pi)\cosh t\,dt
\end{aligned}$$

$$\leqq k^2\sqrt{2}\,\pi^{-1}\,|\,z\,|^{-\frac12}\,|\,2(\nu-\tfrac16)\,|^{-1}\int_{-\infty}^{\infty}\exp(-3k(\operatorname{Re}z^{-1})t^2/\pi+t)\,dt$$

since $-\tfrac12(k+1)\leqq\nu\leqq\tfrac12(k-1)$;

$$=k^2\pi^{-1}(2z)^{-\frac12}\,|\,\nu-\tfrac16\,|^{-1}\exp(\pi/12k(\operatorname{Re}z^{-1}))\int_{-\infty}^{\infty}\exp(-3k(\operatorname{Re}z^{-1})t^2/\pi)\,dt.$$

Now

$$\begin{aligned}k(\operatorname{Re}z^{-1})&=k^2N^{-2}\,|\,z\,|^{-2}\\&\geqq k^2N^{-2}(k^2N^{-4}+k^2/k^2N^2)^{-1}\\&\geqq\tfrac12k^2\geqq\tfrac12.\end{aligned}$$

Hence

$$\begin{aligned}|\,z^{\frac12}J_1^{(\omega)}(\nu;k;z)\,|&\leqq k^2(2\,|\,z\,|)^{-\frac12}\pi^{-1}\,|\,\nu-\tfrac16\,|^{-1}\exp(\pi/6)\\&\qquad\cdot\int_{-\infty}^{\infty}\exp(-6k(\operatorname{Re}z^{-1})t^2/\pi)\sqrt{2}\,dt\\&\leqq k^2\,|\,z\,|^{-\frac12}\sqrt{2}\,|\,\nu-\tfrac16\,|^{-1}\exp(\pi/6)\,(12k(\operatorname{Re}z^{-1}))^{-\frac12}\end{aligned}$$

as in Theorem 2.1,

$$\leqq 2^{-\frac14}3^{-\frac12}\,|\,\nu-\tfrac16\,|^{-1}\exp(\pi/6)N^{\frac12}k.$$

This concludes the proof of Theorem 2.3.

Proof of Theorem 2.4. As before

$$\Omega(q)=\frac{\omega(q)}{P(q)^2}\tag{2.4.1}$$

$$=q^{-\frac34}\sum_{n=0}^{\infty}(-1)^nq^{3(n+\frac12)^2}\frac{1+q^{2n+1}}{1-q^{2n+1}},\ [12;\text{ p. }66],\tag{2.4.2}$$

and if $q=\exp\{\pi i(h+iz)/k\}$,

(2.4.3)
$$q^{\frac34}\Omega(q)=-\tfrac12\sum_{\nu=0}^{2k-1}\sum_{m=-\infty}^{\infty}\frac{(-1)^\nu\exp\{3(\nu+\frac12)^2\pi ih/k-3(2mk+\nu+\frac12)^2\pi z/k\}}{\tanh\{(\nu+\frac12)\pi ih/k-(2mk+\nu+\frac12)\pi z/k\}};$$

by the Poisson summation formula [1; pp. 7-9] we obtain

(2.4.4)
$$\begin{aligned}q^{\frac34}\Omega(q)&=-(2k)^{-1}\sum_{\nu\bmod 2k}(-1)^\nu\exp(3\pi ih\mu^2/k)\\&\qquad\cdot\sum_{n=0}^{\infty}{}^{*}\int_{-\infty}^{\infty}\frac{\exp\{2n\pi i(x-\mu)/2k-3\pi zx^2/k\}}{\tanh(\pi ih\mu/k-\pi zx/k)}\,dx\end{aligned}$$

where $\sum_{n=0}^{\infty}{}^{*}a_n=\tfrac12a_0+\sum_{n=0}^{\infty}a_n$ and $\mu=\nu+\tfrac12$.

Note that if ν is replaced by $\nu + k$, the change introduced is

$$(-1)^k \exp\{3\pi i h(k(\nu+k+1)+k\nu)/k - \pi i n\} = (-1)^{2\nu+2k+1+n}$$
$$= (-1)^{n+1}.$$

Hence if n is even the terms ν and $\nu + k$ cancel each other, and if n is odd then these terms are identical. Thus

$$\begin{aligned} q^{\frac{3}{4}}\Omega(q) &= -k^{-1} \sum_{\nu \bmod k} (-1)^{\nu} \\ &\cdot \exp(3\pi i h\mu^2/k) \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \frac{\exp\{(2n+1)\pi i(x-\mu)/k - 3\pi z x^2/k\}}{\tanh(\pi i h\mu/k - \pi z x/k)}\, dx. \end{aligned} \tag{2.4.4a}$$

As in Theorem 2.3, we define $\mathcal{S}_k(x) = \coth(x/k)\sinh(3x)$, then if

$$\begin{aligned} \mathcal{F}_n(x) &= \frac{\exp\{(2n+1)\pi i(x-\mu)/k - 3\pi z x^2/k\}}{\tanh(\pi i h\mu/k - \pi z x/k)} \\ &= \frac{(-1)^{\nu+\frac{1}{2}(3h-1)} \exp\{(2n+1)\pi i(x-\mu)/k - 3\pi z x^2/k\}\, \mathcal{S}_k(\pi i h\mu - \pi z x)}{i \cosh(3\pi z x)} \end{aligned}$$

$\mathcal{F}_n(x)$ *has* (at most) simple poles at $x_m = (2m+1)i/6z$. The residue at x_m is

$$\lambda_{n,m} =$$
$$\frac{(-1)^{m+\nu+\frac{1}{2}(3h-1)} \exp\{(2n+1)\pi i(x-\mu)/k - 3\pi z x^2/k\}\, \mathcal{S}_k(\pi i h\mu - \pi z x_m)}{3\pi z i},$$

and

$$\lambda_{m+l,m} = \exp\{2\pi i l(x_m - \mu)/k\}\lambda_{m,m}.$$

Now by Cauchy's theorem,

$$\left\{\int_{-\infty}^{\infty} - P\int_{-\infty+x_n}^{\infty+x_n}\right\} \mathcal{F}_n(x)\, dx = 2\pi i\{\lambda_{n,0} + \lambda_{n,1} + \cdots + \tfrac{1}{2}\lambda_{n,n}\}.$$

Hence

$$\begin{aligned} q^{\frac{3}{4}}\Omega(q) &= -2\pi i k^{-1} \sum_{\nu \bmod k} (-1)^{\nu} \\ &\cdot \exp(3\pi i h\mu^2/k) \sum_{n=0}^{\infty} \{\lambda_{n,0} + \lambda_{n,1} + \cdots + \tfrac{1}{2}\lambda_{n,n}\} \\ &- k^{-1} \sum_{\nu \bmod k} (-1)^{\nu} \exp(3\pi i h\mu^2/k) \sum_{n=0}^{\infty} P\int_{-\infty+x_n}^{\infty+x_n} \mathcal{F}_n(x)\ dx \\ &= \Sigma_1 + \Sigma_2. \end{aligned} \tag{2.4.5}$$

Σ_1 yields the main term of the transformation which involves $\omega(q)$ this time, and Σ_2 yields the "error" term.

In Σ_1 we have

$$(2.4.6)\quad \sum_{m=0}^{\infty}\{\lambda_{n,0}+\lambda_{n,1}+\cdots+\tfrac{1}{2}\lambda_{n,n}\}=\tfrac{1}{2}\sum_{m=0}^{\infty}\lambda_{m,m}\frac{1+\exp\{2\pi i(x_m-\mu)/k\}}{1-\exp\{2\pi i(x_m-\mu)/k\}}$$

We now take a closer look at $\lambda_{m,m}$. Since $\lambda_{m,m}$ was a residue of $\mathfrak{F}_m(x)$, we note that $\lambda_{m,m}\neq 0$ if and only if

$$\tanh(\pi ih\mu/\mathrm{k}-\pi zx_m/k)=0,\ \text{i.e.}\ \sin\{\pi(h\mu/k-(2m+1)/6k)=0.$$

Thus we must have

$$3h(2\nu+1)-(2m+1)\equiv 0 \pmod{6k}$$

If $2m+1\not\equiv 0 \pmod 3$, then the above congruence is impossible. Thus $2m+1=6j+3$, and

$$h(2\nu+1)-(2j+1)\equiv 0 \pmod{2k},$$

or

$$2\nu+1\equiv -h''(2j+1) \pmod{2k},$$

or

$$\nu\equiv -\tfrac{1}{2}\{h''(2j+1)+1\} \pmod{k}.$$

Thus when considering $\sum_{\nu \bmod k}$ we see that the only non-zero term is the term $\nu\equiv-\frac{1}{2}\{h''(2j-1)-1\} \pmod{k}$.

We now evaluate $\lambda_{m,m}$ when $m=3j+1$ and $\nu=-\frac{1}{2}\{h''(2j+1)+1\}$.

$$\lambda_{m,m}=\lim_{x\to x_m}\mathfrak{F}_m(x)$$

$$(2.4.7)\qquad =-k(\pi z)^{-1}\exp\{-3\pi(2j+1)^2/4kz+\tfrac{1}{2}3(2j+1)^2\pi ih''/k\}.$$

Hence

$$(2.4.8)\quad \Sigma_1=iz^{-1}(-1)^{\frac{1}{2}(h''+1)}\exp\{-3\pi ih''k''/2+3\pi i(h''+iz^{-1}(/4k\}\cdot\Omega(\exp\{\pi i(h''+iz^{-1})/k\}).$$

We now evaluate Σ_2. In Σ_2 if $n=3p+\delta$, we have

$$(2.4.9)\quad \Sigma_2=-(2k)^{-1}\sum_{\delta=-1}^{1}\sum_{p=-\infty}^{\infty}(-1)^{\nu}\exp\{3\pi ih\mu^2/k-\pi(2n+1)^2/12kz-(2n+1)\pi i\mu/k\}\cdot\int_{-\infty}^{\infty}\frac{\exp(-3\pi zx^2/k)}{\tanh\{\pi ih\mu/k-\pi zx/k-\pi i(6p+2\delta+1)/6k\}}dx.$$

If we replace $(2\nu+1)h-2p$ by $(2\nu+1) \pmod{2k}$, i.e.

$$\nu \to -\tfrac{1}{2}\{h''(2\nu+1+2p)+1\} \pmod{k},$$

then we have $\sum_{p=-\infty}^{\infty}$ free of x and ν. We define

$$U_\delta(z) = \sum_{p=-\infty}^{\infty} (-1)^p \exp\{\pi i h''(3p^2(2\delta+1)p)/k - \pi(6p+2\delta+1)^2/12kz\} \tag{2.4.10}$$

$$= \begin{cases} 0 \text{ if } \delta = 1 \\ \dfrac{\exp(-\pi/12kz)}{P(\exp\{2\pi i(h''+iz^{-1})/k\})} \end{cases} \quad \text{if } \delta = -1 \text{ or } 0, \text{ as before.} \tag{2.4.11}$$

Thus

(2.4.12)
$$\Sigma_2 = \frac{-k^{-1}(-1)^{\frac{1}{2}(h''+1)}\exp(-\pi/12kz - 3\pi i h''k''/2)}{P(\exp\{2\pi i(h''+iz^{-1})/k\})} \sum_{\nu \bmod k} (-1)^\nu$$

$$\cdot \exp(-3\pi i h''\mu^2/k + \pi i h''\mu/k)$$

$$\int_{-\infty}^{\infty} \frac{\exp(-3\pi z x^2/k)}{\tanh(\pi i h\mu/k - \pi z x/k - \pi i/6k)}\, dx.$$

We have established

$$\exp\{3\pi i(h+iz)/4k\} \frac{\omega(\exp\{\pi i(h+iz)/k\})}{P(\exp\{2\pi i(h+iz)/k\})} = \Sigma_1 + \Sigma_2;$$

hence by (2.0.1)

$$\begin{aligned} &\omega(\exp\{\pi i(h+iz)/k\}) \\ &\quad = i(-1)^{\frac{1}{2}(h''+1)}\exp(-3\pi i h''k''/2 + 3\pi i h''/4k - 3\pi i h/4k)\omega_{h,k} \\ &\qquad \cdot z^{-\frac{1}{2}}\exp(-2\pi z^{-1}/3k + 2\pi z/3k)\omega(\exp\{\pi i(h''+iz^{-1})/k\}) \\ &\quad - k^{-1}(-1)^{\frac{1}{2}(h''+1)}\exp(-3\pi i h''k''/2 - 3\pi i h/4k)\omega_{h,k} z^{\frac{1}{2}}\exp(2\pi z/3k) \end{aligned} \tag{2.4.13}$$

$$\cdot \sum_{\nu \bmod k} (-1)^\nu \exp(-3\pi i h''\mu^2/k + \pi i h''\mu/k)$$

$$\cdot \int_{-\infty}^{\infty} \frac{\exp(-3\pi z x^2/k)}{\tanh(\pi i h\mu/k - \pi z x/k - \pi i/6k)}\, dx.$$

Thus when h is odd, the general linear transformation sends ω into ω plus a correction $E_2^{(\omega)}(h,k;z)$; this concludes the proof of Theorem 2.4.

We now proceed to the task of simplifying the expressions for $\epsilon_{h,k}^{(f)}$, $\eta_{h,k}^{(f)}$, $\epsilon_{h,k}^{(\omega)}$, $\eta_{h,k}^{(\omega)}$.

COROLLARY 2.5. *We have*

$$\begin{aligned}
&(2.5.1a)\\
&(2.5.1b)
\end{aligned}\qquad \epsilon_{h,k}^{(f)} = \begin{cases} (-1)^{\frac{1}{2}(k+1)}\omega_{h,2k} & k \text{ odd} \\ i(-1)^{\frac{1}{2}(h-1+k)}\omega_{h,2k} & k \text{ even} \end{cases}$$

$$(2.5.2)\qquad \eta_{h,k}^{(f)} = (-1)^{\frac{1}{2}(k-1)}\exp\{3(k+1)^2\pi ih'/2k\}\omega_{\frac{1}{2}h,k}$$

$$(2.5.3)\qquad \epsilon_{h,k}^{(\omega)} = (-1)^{\frac{1}{2}(k-1)}\exp\{2\pi ih3(k^2-1)/8k\}\omega_{h,k}$$

$$\begin{aligned}
&(2.5.4a)\\
&(2.5.4b)
\end{aligned}\qquad \eta_{h,k}^{(\omega)} = \begin{cases} i(-1)^{\frac{1}{2}(h'''+1)}\exp\{3\pi i(h'''-h)/4k\}\omega_{h,k}, & k \text{ even}, \\ \qquad hh''' \equiv -1 \pmod{8k}; & \\ (-1)^{\frac{1}{2}(k+1)}\exp\{2\pi ih'(5k^2+3)/8k & \\ \qquad +2\pi ih3(k^2-1)/8k\}\omega_{h,k} & k \text{ odd}. \end{cases}$$

Remark. Our results agree with those of Dragonette insofar as is possible to check; however, it appears she has made a slight error in line (7.8) [2; p. 499]. She gives the incorrect evaluation of $\epsilon_{h,k}^{(\omega)}$; if one uses line (7.2) of her paper, one easily derives a result on agreement with ours.

Proof. The proofs of these results involve routine simplifications of the expressions already obtained. For (2.5.1), we use the fact that h is odd, $hh'' = -1-2kk''$, and $\exp(\pi ih''/2) = -\exp(\pi ih/2)$. For (2.5.2), we use the fact that h is even, k is odd, and $k' \equiv -k-khh' \pmod 8$. (2.5.3) is deduced directly. For (2.5.4a), we use the fact that h is odd, k is even, and $h'' \equiv h''' \pmod 4$. For (2.5.4b), we use the fact that h is odd, k is odd, and $k'' \equiv -\frac{1}{2}(-k-khh'') \pmod 8$.

III. The modular transformation formulae of ϕ, ψ, and ν. Having obtained the transformation formulae for f and ω, we obtain the corresponding formulae for ϕ, ψ, and ν by means of certain identities involving the mock theta functions [2; p. 475, (1.2(a)), (1.2(d))].

LEMMA 3.1. *If* $\theta_2(0,q) = \sum_{n=-\infty}^{\infty} q^{(n+\frac{1}{2})^2}$, *and* $\theta_4(0,q) = \sum_{n=-\infty}^{\infty} (-1)^n q^{n^2}$ *with* $hh'+kk' = -1$ *and* g.c.d.$(h,k) = 1$, *then we have for* $(\operatorname{Re} z) > 0$

$$\begin{aligned}\theta_2{}^2(0, \exp\{\pi i(h+iz)/k\} &= iz^{-1}(-1)^{\frac{1}{2}(h-1)}\exp(\pi ihk'/2)\\ &\quad\cdot\theta_2{}^2(0, \exp\{\pi i(h'+iz^{-1})/k\})\\ &\quad h \text{ odd, } k \text{ even, } h' \text{ odd;}\end{aligned}\tag{3.1.1}$$

$$\begin{aligned}\theta_4{}^2(0, \exp\{\pi i(h+iz)/k\} &= z^{-1}(-1)^{\frac{1}{2}(k-1)}\exp(-\pi ih'k/2)\\ &\quad\cdot\theta_2{}^2(0, \exp\{\pi i(h'+iz^{-1})/k\})\\ &\quad h \text{ even, } k \text{ odd, } k' \text{ odd;}\end{aligned}\tag{3.1.2}$$

$$\begin{aligned}\theta_4{}^2(0, \exp\{\pi i(h+iz)/k\}) &= iz^{-1}(-1)^{\frac{1}{2}(h-1)}\exp(-\pi ihk/2)\\ &\quad\cdot\theta_4{}^2(0, \exp\{\pi i(h'+iz^{-1})/k\})\\ &\quad h \text{ odd, } k' \text{ even, } h' \text{ odd;}\end{aligned}\tag{3.1.3}$$

$$\begin{aligned}\theta_2{}^2(0, \exp\{\pi i(h+iz)/k\}) &= z^{-1}(-1)^{\frac{1}{2}(k-1)}\exp(\pi ihk/2)\\ &\quad\cdot\theta_4{}^2(0, \exp\{\pi i(h'+iz^{-1})/k\})\\ &\quad k \text{ odd, } k' \text{ odd, } h' \text{ even;}\end{aligned}\tag{3.1.4}$$

Proof. These formulae are obtained directly from [11] using Table XX on page 241 and Table XLII on page 262. In Table XX, we let $-h=a$, $k=b$, $k'=c$, $h'=d$, $\tau=(h+iz)/k$, $T=(h'+iz^{-1})/k$. We obtain

(3.1.1) with k' even from line 1 of Table XX, and lines (2), (6), (7), (8) of XLII;

(3.1.1) with k' odd from line 2 of Table XX, and lines (2), (6), (7), (8) of XLII;

(3.1.2) with h' odd from line 6 of Table XX, and lines (2), (6), (7), (8) of XLII;

(3.1.2) with h' even from line 5 of Table XX, and lines (2), (6), (7), (8) of XLII;

(3.1.3) with k even from line 1 of Table XX, and lines (4), (6), (7), (8) of XLII;

(3.1.3) with k odd from line 3 of Table XX, and lines (4), (6), (7), (8) of XLII;

(3.1.4) with h odd from line 4 of Table XX, and lines (4), (6), (7), (8) of XLII;

(3.1.4) with h even from line 5 of Table XX, and lines (4), (6), (7), (8) of XLII.

In the next two lemmas, we study two theta functions closely related to the mock theta functions.

LEMMA 3.2. *If* $G_1(q)=\theta_4(0,q)\prod_{r=1}^{\infty}(1+q^r)^{-1}$ *and*

$$G_2(q)=\tfrac{1}{2}q^{-\frac{1}{4}}\theta_2(0,q)\prod_{r=1}^{\infty}(1+q^{2r}),$$

then

$$G_1(\exp\{\pi i(h+iz)/k\})=\begin{cases}2^{\frac{3}{2}}z^{-\frac{1}{2}}(-1)^{h'}\exp(-\pi ih'/k)\eta_{h,k}^{(f)}\exp(-\pi/3kz \\ \qquad -\pi z/24k)G_2(\exp\{\pi i(h'+iz^{-1})/k\}) \quad h \text{ even} \\ z^{-\frac{1}{2}}(-1)^k\epsilon_{h,k}^{(f)}\exp(\pi/24kz \\ \qquad -\pi z/24k)G_1(\exp\{\pi i(h''+iz^{-1})/k\}) \quad h \text{ odd}\end{cases}$$

where in the top line $hh'\equiv -1 \pmod{k}$ and in the bottom line $hh''\equiv -1 \pmod{2k}$. $\epsilon_{h,k}^{(f)}$ appears in Theorem 2.2, and $\eta_{h,k}^{(f)}$ appears in Theorem 2.1.

Proof. We note that

$$\begin{aligned}G_1(q)&=\theta_4(0,q)\prod_{r=1}^{\infty}(1+q^r)^{-1}\\&=\theta_4(0,q)\prod_{r=1}^{\infty}(1+q^r)^{-1}\prod_{r=1}^{\infty}(1-q^r)\prod_{r=1}^{\infty}(1-q^r)^{-1}\\&=\theta_4^2(0,q)P(q) \qquad [5;\ \text{p. } 185]\end{aligned}$$

$$\begin{aligned}G_2(q)&=\tfrac{1}{2}q^{-\frac{1}{4}}\theta_2(0,q)\prod_{r=1}^{\infty}(1+q^{2r})\\&=\tfrac{1}{2}q^{-\frac{1}{4}}\theta_2(0,q)\prod_{r=1}^{\infty}(1+q^{2r})\prod_{r=1}^{\infty}(1-q^{4r})\prod_{r=1}^{\infty}(1-q^{4r})^{-1}\\&=\tfrac{1}{4}q^{-\frac{1}{2}}\theta_2^2(0,q)P(q^4) \qquad [5;\ \text{p. } 185].\end{aligned}$$

The remainder of the proof is routine. Since $G_1(q)=\theta_4^2(0,q)P(q)$, we apply Lemma 3.1 to $\theta_4^2(0,q)$, and (2.0.1) to $P(q)$; the desired result follows directly using the expressions for $\epsilon_{h,k}^{(f)}$ and $\eta_{h,k}^{(f)}$ found in Corollary 2.5.

LEMMA 3.3. *If* G_1 *and* G_2 *are as in the previous lemma, then*

$$G_2(\exp\{\pi i(h+iz)/k\}) = \begin{cases} (2z)^{-\frac{1}{2}}(-1)^h \exp(\pi i h/k)\epsilon_{2h,k}^{(\omega)}\exp(\pi z/3k \\ \qquad + \pi z^{-1}/24k)G_1[(-1)^{h'}\exp\{\pi i(h'+iz^{-1})/k\}] \\ \qquad\qquad k \text{ odd} \\ z^{-\frac{1}{2}}(-1)^{k'}\exp\{\pi i(h-h')/k\}\eta_{h,\frac{1}{2}k}^{(\omega)}\exp(\pi z/3k \\ \qquad - \pi z^{-1}/3k)G_2(\exp\{\pi i(h'+iz^{-1})/k\}) \\ \qquad\qquad k \text{ even} \end{cases}$$

where $hh' + kk' = -1$.

Proof. The proof of this lemma is exactly like that of Lemma 3.2. Since $G_2(q) = \frac{1}{4}q^{-\frac{1}{2}}\theta_2{}^2(0,q)P(q^4)$, we apply Lemma 3.1 to $\theta_2{}^2(0,q)$ and (2.0.1) to $P(q^4)$; the desired result follows directly using the expressions $\epsilon_{h,k}^{(\omega)}$ and $\eta_{h,k}^{(\omega)}$ found in Corollary 2.5.

To obtain the linear transformation formulae for ϕ, ψ, and v we utilize the following identities partially stated by Ramanujan [9; p. 354] and proved by Watson [12; p. 63].

(A) $$2\phi(q) = f(-q) + G_1(-q)$$

(B) $$4\psi(q) = -f(-q) + G_1(-q)$$

(C) $$\mp q^{-1}v(\pm q) = \omega(q^2) \mp q^{-1}G_2(q)$$

where G_1 and G_2 are given in Lemma 3.2,

THEOREM 3.4. *Let* g. c. d. $(h,k) = 1$, $hh' + kk' = -1$, *also if h is odd* $hh'' + 2kk'' = -1$, *and* $\operatorname{Re} z > 0$ *with the principal branch of* $z^{\frac{1}{2}}$ *taken. Then*

$$\begin{array}{ll} (3.4.1) & \\ (3.4.2) & \phi[\exp\{\pi i(h+iz)/k\}] = \\ (3.4.3) & \end{array} \begin{cases} z^{-\frac{1}{2}}\epsilon_{h,k}^{(\phi)}\exp(-\pi z/24k + \pi z^{-1}/24k)\ . \\ \qquad \cdot\phi(\exp\{\pi i(h''+iz^{-1})/k\}) \\ \qquad\quad + \frac{1}{2}E_2^{(f)}(h+k,k;z) \quad (k \text{ even}) \\ 2^{\frac{1}{2}}z^{-\frac{1}{2}}\eta_{h,k}^{(\phi)}\exp(-\pi z/24\mathrm{k} - \pi z^{-1}/3k) \\ \qquad \cdot v[(-1)^{h'-1}\exp\{\pi i(h'+iz^{-1})/k\}] \\ \qquad\quad + \frac{1}{2}E_1^{(f)}(h+k,k;z) \quad (k \text{ odd, } h \text{ odd}) \\ 2z^{-\frac{1}{2}}\sigma_{h,k}^{(\phi)}\exp(-\pi z/24k + \pi z^{-1}/24k) \\ \qquad \cdot\psi[(-1)^{h'-1}\exp\{\pi i(h'+iz^{-1})/k\}] \\ \qquad\quad + \frac{1}{2}E_2^{(f)}(h+k,k;z) \quad (h \text{ even}) \end{cases}$$

$$\begin{array}{ll} (3.4.4) & \\ (3.4.5) & \psi[\exp\{\pi i(h+iz)/k\}] = \\ (3.4.6) & \end{array} \begin{cases} \frac{1}{2}z^{-\frac{1}{2}}\epsilon_{h,k}^{(\psi)}\exp(-\pi z/24k + \pi z^{-1}/24k) \\ \qquad \cdot\phi[(-1)^{h'}\exp\{\pi i(h'+iz^{-1})/k\}] \\ \qquad\quad - \frac{1}{4}E_2^{(f)}(h+k,k;z) \quad (h \text{ even}) \\ (2z)^{-\frac{1}{2}}\eta_{h,k}^{(\psi)}\exp(-\pi z/24k - \pi z^{-1}/3k) \\ \qquad \cdot v[(-1)^{h'}\exp\{\pi i(h'+iz^{-1})/k\}] \\ \qquad\quad - \frac{1}{4}E_1^{(f)}(h+k,k;z) \quad (h \text{ odd, } k \text{ odd}) \\ z^{-\frac{1}{2}}\sigma_{h,k}^{(\psi)}\exp(-\pi z/24k + \pi z^{-1}/24k) \\ \qquad \cdot\psi(\exp\{\pi i(h''+iz^{-1})/k\}) \\ \qquad\quad - \frac{1}{4}E_2^{(f)}(h+k,k;z) \quad (k \text{ even}) \end{cases}$$

$$v(\exp\{\pi i(h+iz)/k\}) = \begin{cases} (2z)^{-\frac{1}{2}}\epsilon_{h,k}{}^{(v)}\exp(\pi z/3k+\pi z^{-1}/3k) & (3.4.7)\\ \quad\cdot\phi[(-1)^{h'-1}\exp\{\pi i(h'+iz^{-1})/k\}] & \\ \quad\quad -\exp\{\pi i(h+iz)/k\}E_1{}^{(\omega)}(2h,k;2z) & \\ \quad\quad\quad (h \text{ odd, k odd}) & \\[2ex] z^{-\frac{1}{2}}\eta_{h,k}{}^{(v)}\exp(\pi z/3k-\pi z^{-1}/3k) & (3.4.8)\\ \quad\cdot v[(-1)^{k'}\exp\{\pi i(h'+iz^{-1})/k\}] & \\ \quad\quad -\exp\{\pi i(h+iz)/k\}E_2{}^{(\omega)}(h,\tfrac{1}{2}k;z) & \\ \quad\quad\quad (k \text{ even}) & \\[2ex] 2^{\frac{1}{2}}z^{-\frac{1}{2}}\sigma_{h,k}{}^{(v)}\exp(\pi z/3k+\pi z^{-1}/24k) & (3.4.9)\\ \quad\cdot\psi[(-1)^{h'-1}\exp\{\pi i(h'+iz^{-1})k] & \\ \quad -\exp\{\pi i(h+iz)/k\}E_1{}^{(\omega)}(2h+2k,k;2z) & \\ \quad\quad\quad (h \text{ even}) & \end{cases}$$

with

$$\epsilon_{h,k}{}^{(\phi)}=\epsilon_{h+k,k}{}^{(f)} \quad (k \text{ even})$$

$$\eta_{h,k}{}^{(\phi)}=(-1)^{h'}\exp(-\pi ih'/k)\eta_{h+k,k}{}^{(f)} \quad (h+k \text{ even})$$

$$\sigma_{h,k}{}^{(\phi)}=-\epsilon_{h+k,k}{}^{(f)} \quad (h \text{ even})$$

$$\epsilon_{h,k}{}^{(\psi)}=-\epsilon_{h+k,k}{}^{(f)} \quad (h \text{ even})$$

$$\eta_{h,k}{}^{(\psi)}=(-1)^{h'}\exp(-\pi ih'/k)\eta_{h+k,k}{}^{(f)} \quad (h+k \text{ even})$$

$$\sigma_{h,k}{}^{(\psi)}=\epsilon_{h+k,k}{}^{(f)} \quad (k \text{ even})$$

$$\epsilon_{h;k}{}^{(v)}=-\exp(\pi ih/k)\epsilon_{2h,k}{}^{(\omega)} \quad (h+k \text{ even})$$

$$\eta_{h,k}{}^{(v)}=(-1)^{k'}\exp\{\pi i(h-h')/k\}\eta_{h,\frac{1}{2}k}{}^{(\omega)} \quad (k \text{ even})$$

$$\sigma_{h,k}{}^{(v)}=\exp(\pi ih/k)\epsilon_{2h,k}{}^{(\omega)} \quad (h \text{ even})$$

where $\epsilon_{h,k}{}^{(f)}$, $\eta_{h,k}{}^{(f)}$, $\epsilon_{h,k}{}^{(\omega)}$, $\eta_{h,k}{}^{(\omega)}$ *and* $E_1{}^{(f)}(h,k;z)$, $E_2{}^{(f)}(h,k;z)$, $E_1{}^{(\omega)}(h,k;z)$, $E_2{}^{(\omega)}(h,k;z)$ *are given in Theorems* 2.1, 2.2, 2.3 *and* 2.4 *respectively.*

Proof. We shall derive (3.4.1) only. The remaining formulae are proved in a similar manner.

$$\begin{aligned} (3.4.10)\quad & 2\phi(\exp\{\pi i(h+iz)/k\}) \\ & =f(\exp\{\pi i[(h+k)+iz]/k\}+G_1(\exp\{\pi i[(h+k)+iz]/k\}). \end{aligned}$$

Now for k even, we use the fact that $(h+k)''\equiv h''+k \pmod{2k}$ since $(h+k)(h''+k)=hh''+k(h+h''+k)\equiv-1 \pmod{2k}$. Hence by (A)

$$\begin{aligned} (3.4.11)\quad & 2\phi(-\exp\{\pi i[(h+k)''+iz^{-1}]/k\}) \\ & =f(\exp\{\pi i[(h+k)''+iz^{-1}]/k\})+G_1(\exp\{\pi i[(h+k)''+iz^{-1}]/k\}). \end{aligned}$$

Multiplying (3.4.11) by $z^{-\frac{1}{2}}\epsilon_{h+k,k}^{(f)}\exp(-\pi z/24k+\pi z^{-1}/24k)$, then subtracting from (3.4.10) and using Theorem 2.2 and Lemma 3.2, we get

$$\begin{aligned}&\phi(\exp\{\pi i(h+iz)/k\})\\ &=z^{-\frac{1}{2}}\epsilon_{h+k,k}^{(f)}\exp(-\pi z/24k+\pi z^{-1}/24k)\\ &\quad\cdot\phi(\exp\{\pi i(h''+iz^{-1})/k\})+\tfrac{1}{2}E_2^{(f)}(h+k,k;z)\end{aligned}\tag{3.4.12}$$

which is (3.4.1).

IV. The exponential sums. Before evaluating the exponential sums that arise for the mock theta functions, we give the following lemma which will be useful in the next section.

LEMMA 4.1. *If* g.c.d.$(D,k)=1$, $d\mid 24$, *and* $0\leqq\sigma_1<\sigma_2\leqq k$,

$$A_k(D;d;F,G;\sigma_1;\sigma_2)$$

$$=\sum_{\substack{h \bmod dk\\ \sigma_1\leqq Dh^*\leqq\sigma_2\\ \text{g.c.d.}(h,dk)=1}}\omega_{h,k}\exp(-2\pi iFh/dk\ 2\pi iGh^*/dk),$$

where $\omega_{h,k}$ *appears in* (2.0.1), *and* $hh^*\equiv-1 \pmod{dk}$, *then*

$A_k(D;d;F,G;\sigma_1;\sigma_2)=O(((24F+d)/d,k)^{\frac{1}{3}}k^{\frac{2}{3}+\epsilon})$, *as* $k\to+\infty$, *where the constant implied by the O-symbol is absolute.*

Proof. Many papers contain theorems like this for sums related to the standard Kloosterman sum (e.g. [3], [7]). However, by [9; p. 284] we have

$$\omega_{h,k}=-i\left(\frac{-k}{h}\right)\exp\{2\pi i[(4k^2+3k-1)h-(k^2-1)\tilde{h}]/24k\}$$

if k is even, g.c.d.$(h,24)=\Theta$ and $h\tilde{h}\equiv-1\ (\bmod 24k/\Theta)$;

$$\omega_{h,k}=\exp\{-\pi i(k-1)/4\}\left(\frac{-h}{k}\right)\exp\{2\pi i[(k^2-1)h-(k^2-1)\tilde{h}]/24k\}$$

if k is odd, g.c.d.$(h,24)=\Theta$ and $h\tilde{h}\equiv-1\ (\bmod 24k/\Theta)$. Thus we see that the sum we are considering is in fact related to a generalized Kloosterman sum; such sums have recently been studied in detail in [13]. A generalized Klosterman sum is of the form

$$\sum_{\substack{h \bmod k \\ \text{g.c.d.}(h,k)=1}} \chi(h)\exp\{2\pi i(ah-h')/k\}$$

where $\chi(h)$ is a quadratic character. This sum satisfies the same estimates as the original Kloosterman sum; one may easily prove that such sums are multiplicative, and one has

$$\Big| \sum_{\substack{h \bmod p^\alpha \\ \text{g.c.d.}(h,p)=1}} \chi(h)\exp\{2\pi i(ah-h')/p^\alpha\} \Big| \leqq 2p^{\frac{1}{2}\alpha}, \qquad \alpha > 1$$

by [6; p. 283] and [10; p. 102, line (54)].

This information gives by the method of Lehner [7; § 10]

$$| A_k(D;d;F,G;\sigma_1,\sigma_2) = \begin{cases} O\{(-24F/d+4k^2+3k-1, 24k)^{\frac{1}{2}}k^{\frac{1}{2}+\epsilon}\} & (k \text{ even}) \\ O\{(-24F/d+k^2-1, 24k)^{\frac{1}{2}}k^{\frac{1}{2}+\epsilon}\} & (k \text{ odd}) \end{cases}$$
$$= O\{(24F/d+1, k)^{\frac{1}{2}}k^{\frac{1}{2}+\epsilon}\}.$$

We now go to the exponential sums in question. The following equations are direct consequences of Corolary 2. 5. We shall only exhibit the exponential sums necessary to deal with Theorem 5. 1. Examination of Theorem 3. 4 shows that analogous results hold for the exponential sums relevant in the proof of Theorem 5. 2. If we let

$$\lambda_k^{(f)}(\epsilon;\sigma_1;\sigma_2;n) = \sum_{\substack{h \bmod 2k \\ \text{g.c.d.}(h,2k)=1 \\ \sigma_1 \leqq h' \leqq \sigma_2}} \epsilon_{h,k}^{(f)}\exp(-\pi ihn/k), \tag{4.2}$$

then

$$\lambda_k^{(f)}(\epsilon;\sigma_1;\sigma_2;n) = \begin{cases} (-1)^{\frac{1}{2}(k+1)}\sum_{j=0}^{1} A_{2k}(1;1;n,0;\sigma_1-jk,\sigma_2-jk) & (k \text{ odd}) \\ (-1)^{\frac{1}{2}k}\sum_{j=0}^{1} \mathrm{A}_{2k}(1,1;n-\frac{1}{2}k,0;\sigma_1-jk,\sigma_2-jk) & (k \text{ even}) \end{cases}$$

If we let

$$\lambda_k^{(f)}(\eta;\sigma_1;\sigma_2;n) = \sum_{\substack{h \bmod 2k \\ \text{g.c.d.}(h,k)=1 \\ 2 \mid h \\ \sigma_1 \leqq h' \leqq \sigma_2}} \eta_{h,k}^{(f)}\exp(-\pi ihn/k), \tag{4.3}$$

then

$$\lambda_k^{(f)}(\eta;\sigma_1;\sigma_2;n) = (-1)^{\frac{1}{2}(k-1)}A_k(\tfrac{1}{2}(k+1);1;n,3(k+1)^3/8;\sigma_1;\sigma_2).$$

The following two exponential sums will also arise when we estimate the error terms in the asymptotic expansions. If we let

$$(4.4)\quad \Lambda_{k,1}^{(f)}(\delta;\sigma_1,\sigma_2;n)=\sum_{\substack{h \bmod 2k\\ \text{g.c.d.}(h,2k)=1\\ 2\mid h\\ \sigma_1\leqq h'\leqq\sigma_2\\ \sigma_1\leqq Dh^*\leqq\sigma_2}}\delta_1^{(f)}(\nu;\tfrac{1}{2}h;2h';k)\exp(-\ ihn/k),$$

then

$$\Lambda_{k,1}^{(f)}(\delta;\sigma_1,\sigma_2;n)=(-1)^{\nu}\sum_{\substack{h \bmod k\\ \text{g.c.d.}(2h,k)=1\\ \sigma_1\leqq\frac{1}{2}(k+1)h'\leqq\sigma_2}}\delta_1^{(f)}(\nu;h;h';k)\exp(-2\pi ihn/k)$$

$$=(-1)^{\nu}A_k(\tfrac{1}{2}(k+1);1;n,\tfrac{1}{2}(\nu-3\nu^2);\sigma_1,\sigma_2).$$

If we let

$$(4.5)\quad \Lambda_{k,2}^{(f)}(\delta;\sigma_1,\sigma_2;n)=\sum_{\substack{h \bmod 2k\\ \text{g.c.d.}(h,2k)=1\\ \sigma_1\leqq h'\leqq\sigma_2}}\delta_1^{(f)}(\nu;h;h'';2k)\exp(-\pi ihn/k),$$

then

$$\Lambda_{k,2}^{(f)}(\delta;\sigma_1,\sigma_2;n)$$
$$=(-1)^{\nu}\sum_{j=0}^{1}A_{2k}(1;1;n,\tfrac{1}{2}(\nu-3\nu^2);\sigma_1-jk;\sigma_2-jk).$$

We now treat the exponential sums which actually appear in the asymptotic expansions.

THEOREM 4.6. *If*

$$(4.6.1)\quad \lambda(k)=\tfrac{1}{2}\sum_{\substack{h \bmod 2k\\ \text{g.c.d.}(h,2k)=1}}\epsilon_{h,k}^{(f)}\exp(-\pi ihn/k)\,;$$

$$(4.6.2)\quad \lambda^{(\phi)}(k)=\tfrac{1}{2}\sum_{\substack{h \bmod 2k\\ \text{g.c.d.}(h,2k)=1\\ 2\mid k}}\epsilon_{h,k}^{(\phi)}\exp(-\pi ihn/k)\,;$$

$$(4.6.3)\quad \lambda^{(\psi)}(k)=\tfrac{1}{2}\sum_{\substack{h \bmod 2k\\ \text{g.c.d.}(h,k)=1\\ 2\mid h}}\epsilon_{h,k}^{(\psi)}\exp(-\pi ihn/k)\,;$$

$$(4.6.4)\quad \lambda^{(\nu)}(k)=\tfrac{1}{2}\sum_{\substack{h \bmod 2k\\ \text{g.c.d.}(h,k)=1\\ 2\nmid h,2\nmid k}}\epsilon_{h,k}^{(\nu)}\exp(-\pi ihn/k)\,;$$

$$(4.6.5)\qquad \lambda^{(\omega)}(k)=\tfrac{1}{2}\sum_{\substack{h \bmod 2k\\ \text{g.c.d.}(h,k)=1\\ 2\mid h}} \epsilon_{h,k}{}^{(\omega)}\exp(-\pi i h n/k)\text{ ; } \textit{then}$$

$$(4.6.6)\qquad \lambda(k)=\begin{cases}\tfrac{1}{2}(-1)^{\frac{1}{2}(k+1)}A_{2k}(n) & (k \text{ odd})\\ \tfrac{1}{2}(-1)^{\frac{1}{2}k}A_{2k}(n-\tfrac{1}{2}k) & (k \text{ even}),\end{cases}$$

$$(4.6.7)\qquad \lambda^{(\phi)}(k)=\tfrac{1}{2}(-1)^{n+\frac{1}{2}k}A_{2k}(n-\tfrac{1}{2}k)$$

$$(4.6.8)\qquad \lambda^{(\psi)}(k)=\tfrac{1}{2}(-1)^{n+\frac{1}{2}(k-1)}A_{2k}(n)$$

$$(4.6.9)\qquad \lambda^{(\nu)}(k)=\tfrac{1}{2}(-1)^{n+\frac{1}{2}(k-1)}A_{k}\{(n-1)[\tfrac{1}{2}(k+1)]^{2}-3(k^{2}-1)/8\}$$

$$(4.6.10)\qquad \lambda^{(\omega)}(k)=\tfrac{1}{2}(-1)^{\frac{1}{2}(k-1)}A_{k}\{\tfrac{1}{2}n(k+1)-3(k^{2}-1)/8\}$$

where $A_k(n)=\sum_{\substack{h \bmod k\\ \text{g.c.d.}(h,k)=1}} \omega_{h,k}\exp(-2\pi i h n/k)$ *and* $\omega_{h,k}$ *appears in* (2.0.1).

Proof. (4.6.6) is a special case of (4.2). (4.6.7) follows immediately from the definition of $\epsilon_{h,k}{}^{(\phi)}$ in Theorem 3.4. (4.6.8) also follows immediately from the definition of $\epsilon_{h,k}{}^{(\psi)}$ in Theorem 3.4. We deal with (4.6.10) next. We have

$$\begin{aligned}\lambda^{(\omega)}(k)&=\tfrac{1}{2}\sum_{\substack{h \bmod 2k\\ \text{g.c.d.}(h,k)=1\\ 2\mid h}} \epsilon_{h,k}{}^{(\omega)}\exp(-\pi i h n/k)\\
&=\tfrac{1}{4}(-1)^{\frac{1}{2}(k-1)}\sum_{j=0}^{1}\sum_{\substack{h \bmod 2k\\ \text{g.c.d.}(h,k)=1}} \omega_{h,k}\exp\{\pi i h j+2\pi i h 3(k^{2}-1)/8k-\pi i h n/k\}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\text{[by Corollary 2.5]}\\
&=\tfrac{1}{4}(-1)^{\frac{1}{2}(k-1)}\sum_{j=0}^{1}\sum_{\substack{h \bmod k\\ \text{g.c.d.}(h,k)=1}} \omega_{h,k}\exp\{\pi i h j+2\pi i h 3(k^{2}-1)/8k-\pi i h n/k\}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\cdot(1+(-1)^{jk+n})\\
&=\tfrac{1}{2}(-1)^{\frac{1}{2}(k-1)}\sum_{\substack{h \bmod k\\ \text{g.c.d.}(h,k)=1}} \omega_{h,k}\exp(-2\pi i h\{\tfrac{1}{2}n(k+1)-3(k^{2}-1/8\}/k)\\
&=\tfrac{1}{2}(-1)^{\frac{1}{2}(k-1)}A_{k}\{\tfrac{1}{2}n(k+1)-3(k^{2}-1)/8\}.\end{aligned}$$

$$\begin{aligned}\lambda^{(\nu)}(k)&=-\tfrac{1}{2}\sum_{\substack{h \bmod 2k\\ \text{g.c.d.}(h,k)=1\\ 2\nmid h,\,2\nmid k}} \epsilon_{2h,k}{}^{(\omega)}\exp(\pi i h/k-\pi i h n/k)\qquad\text{[by Theorem 3.4]}\\
&=\tfrac{1}{2}(-1)^{n}\sum_{\substack{h \bmod 2k\\ \text{g.c.d.}(2h,k)=1}} \epsilon_{2h,k}{}^{(\omega)}\exp\{-\pi i h(n-1)/k\}\end{aligned}$$

$$= \tfrac{1}{4}(-1)^n \sum_{j=0}^{1} \sum_{\substack{h \bmod 2k \\ \text{g.c.d.}(h,k)=1}} \epsilon_{2h,k}{}^{(\omega)} \exp(-\pi i h(n-1)/k + \pi i h j)$$

$$= \tfrac{1}{2}(-1)^n \sum_{\substack{h \bmod k \\ \text{g.c.d.}(h,k)=1}} \epsilon_{2h,k}{}^{(\omega)} \exp\{-\pi i h(n-1)(k+1)/k\}$$

$$= \tfrac{1}{2}(-1)^n \sum_{\substack{h \bmod 2k \\ \text{g.c.d.}(h,k)=1 \\ 2 \mid h}} \epsilon_{h,k}{}^{(\omega)} \exp\{-\pi i h(n-1)\tfrac{1}{2}(k+1)\}$$

$$= \tfrac{1}{2}(-1)^{n+\frac{1}{2}(k-1)} A_k\{(n-1)[\tfrac{1}{2}(k+1)]^2 - 3(k^2-1)/8\} \text{ by } (4.6.10).$$

This concludes Theorem 4.6.

V. The asymptotic formulae. In [2], Dragonette has given asymptotic expansions for the power series coefficients of f, ω, ϕ, ψ, and ν (as discussed in § I). In this section we shall improve her results.

Since Dragoneete obtained the main terms for the expansion of $A(n)$ using the Hardy-Ramanujan-Rademacher method, we shall refer the reader to her paper for many of the details. Our method centers upon the use of estimates of generalized Kloosterman sums together with our precise knowledge of the error terms involved in the modular transformation formulae. We now prove the theorem stated in the introduction.

Proof of Theorem 5.1. For the initial Farey dissection we refer the reader to [2; pp. 489-90]. Dragonette obtains there

$$(5.1.1) \quad A(n) = \tfrac{1}{2}\sum_{k=1}^{N} \sum_{\substack{h \bmod 2k \\ \text{g.c.d.}(h,2k)=1}} \exp(-\pi i h n/k) \int_{-\theta'_{h,k}}^{\theta''_{h,k}} \{z^{-\frac{1}{2}}\epsilon_{h,k}{}^{(f)}$$

$$\cdot \exp(\pi z n/k - \pi z/24k + \pi z^{-1}/24k) f(\exp\{\pi i(h'' + iz^{-1})/k\})$$
$$+ E_2{}^{(f)}(h,k;z)\} d\phi$$
$$+ \tfrac{1}{2}\sum_{k=1}^{N} \sum_{\substack{h \bmod 2k \\ \text{g.c.d.}(h,k)=1 \\ 2 \mid h}} \exp(-\pi i h n/k) \int_{-\theta'_{h,k}}^{\theta''_{h,k}} \{2^{\frac{3}{2}} z^{-\frac{1}{2}} \eta_{h,k}{}^{(f)}$$

$$\cdot \exp(\pi z n/k - \pi z/24k - 4\pi z^{-1}/3k)\, \omega(\exp\{2\pi i(h' + iz^{-1})/k\})$$
$$+ E_1{}^{(f)}(h,k;z)\} d\phi$$

$$(5.1.2) \qquad = \sum_{k=1}^{N} \frac{\lambda(k)\exp\{\pi(n-1/24)^{\frac{1}{2}}/k6^{\frac{1}{2}}\}}{k^{\frac{1}{2}}(n-1/24)^{\frac{1}{2}}} + O(\sum_{k=1}^{N}(n,k)^{\frac{1}{2}}k^{\epsilon-1}) + O(n^{\epsilon})$$

$$+\tfrac{1}{2}\sum_{k=1}^{N}\sum_{\substack{h \bmod 2k\\ \text{g.c.d.}(h,2k)=1}}\exp(-\pi ihn/k)\int_{-\theta'_{h,k}}^{\theta''_{h,k}}E_2^{(f)}(h,k;z)d\phi$$

$$+\tfrac{1}{2}\sum_{k=1}^{N}\sum_{\substack{h \bmod 2k\\ \text{g.c.d.}(h,2k)=1}}\exp(-\pi ihn/k)\int_{-\theta'_{h,k}}^{\theta''_{h,k}}E_1^{(f)}(h,k;z)d\phi$$

$$(5.1.3) \qquad = \Sigma_1 + O(n^{\epsilon}) + \Sigma_2 + \Sigma_3,$$

We remark that (5.1.3) is obtained from (5.1.1) by almost the same method that Dragonette uses in obtaining [2; p. 493, line (5.5)] from [2; p. 490, line (4.5)]. The main alteration in the proof is our use of estimates for truncated generalized Kloosterman sums given by Lemma 4.1 and equations (4.2) and (4.3); this is done in exactly the way that Rademacher modifies the Hardy-Ramanujan method to obtain Fourier coefficients for the modular invariant, $J(\tau)$, [8]. Also we exhibit the contributions of $E_1^{(f)}(h,k;z)$ and $E_2^{(f)}(h,k;z)$ whereas Dragonette uses the estimate

$$|E_j^{(f)}(h,k;z)| = O(k\log k) \qquad j=1,2.$$

We shall show in fact that

$$|\Sigma_j| = O(n^{\epsilon}) \qquad j=2,3.$$

For reference we give the following resumè of the notation used above.

$$\theta'_{h,k} = k^{-1}(k_1+k)^{-1}, \qquad \theta''_{h,k} = k^{-1}(k_2+k)^{-1}$$

where $h_1/k_1 < h/k < h_2/k_2$ are adjacent Farey fractions in the Farey series of order N. As a simple consequence of the theory of Farey series we have

$$k_1 \equiv -h' \pmod{k}, \qquad k_2 \equiv h' \pmod{k}, \text{ and}$$
$$N-k < k_1 \leqq N, \qquad N-k < k_2 \leqq N.$$
$$z = kN^{-2} - ik\phi.$$

Also throughout our work here we take $N=[n^{\frac{1}{2}}]$. Using the notation of Theorems 2.1 and 2.2, we have

$$\Sigma_2 = \tfrac{1}{2}\sum_{k=1}^{N}k^{-1}\sum_{\substack{h \bmod 2k\\ \text{g.c.d.}(h,2k)=1}}\exp(-\pi ihn/k)\sum_{\nu \bmod k}\delta_1^{(f)}(\nu;h;h'';2k)$$
$$\cdot\int_{-\theta'_{h,k}}^{\theta''_{h,k}}z^{\frac{1}{2}}J_1^{(f)}(\nu;2k;4z)\,d\phi$$

$$= \tfrac{1}{2}\sum_{k-1}^{N} k^{-1} \sum_{\nu=1}^{k} \sum_{\substack{h \bmod 2k \\ \text{g.c.d.}(h,2k)=1}} \delta_1^{(f)}(\nu; h; h''; 2k) \exp(-\pi i h n/k)$$

$$\cdot \left\{ \int_{-k^{-1}(N+k)^{-1}}^{k^{-1}(N+k)^{-1}} + \int_{-k^{-1}(k_1+k)^{-1}}^{-k^{-1}(N+k)^{-1}} + \int_{k^{-1}(N+k)^{-1}}^{k^{-1}(k_2+k)^{-1}} \right\}$$

$$= \Sigma_{21} + \Sigma_{22} + \Sigma_{23}.$$

$$| \Sigma_{21} | = O\Big(\sum_{k=1}^{N} k^{-1} (n,k)^{\frac{1}{2}} k^{\frac{1}{2}+\epsilon'} (kN)^{-1} \sum_{\nu=1}^{k} N^{\frac{1}{2}} k \,|\, k - 2(\nu - \tfrac{1}{6}) |^{-1} \Big)$$

(by Theorem 2.1 and (4.5))

$$= O\Big(N^{-\frac{1}{2}} \sum_{k=1}^{N} (n,k)^{\frac{1}{2}} k^{-\frac{1}{2}-2\epsilon''}\Big)$$

$$= O\Big(N^{-\frac{1}{2}} \sum_{\substack{d \mid n \\ d \leqq n^{\frac{1}{2}}}} d^{\frac{1}{2}} \sum_{\substack{k=1 \\ \text{g.c.d.}(n,k)=1}}^{n^{\frac{1}{2}}} k^{-\frac{1}{2}+2\epsilon''}\Big)$$

$$= O\Big(N^{-\frac{1}{2}} \sum_{d \mid n} d^{\frac{1}{2}} \sum_{\substack{k'=1 \\ d \leqq n^{\frac{1}{2}}}}^{n^{\frac{1}{2}}/d} d^{-\frac{1}{2}+2\epsilon''} k'^{-\frac{1}{2}+2\epsilon''}\Big)$$

$$= O\Big(N^{-\frac{1}{2}} \sum_{\substack{d \mid n \\ d \leqq n^{\frac{1}{2}}}} d^{2\epsilon''} n^{\frac{1}{4}+\epsilon''} d^{-\frac{1}{2}-2\epsilon''}\Big)$$

$$= O\Big(n^{\epsilon''} \sum_{\substack{d \mid n \\ d \leqq n^{\frac{1}{2}}}} d^{-\frac{1}{2}}\Big)$$

$$= O(n^{\epsilon}).$$

Σ_{22} and Σ_{23} are of the same structure so we treat only Σ_{23}.

$$\Sigma_{23} = \tfrac{1}{2}\sum_{k=1}^{N} k^{-1} \sum_{\nu=1}^{k} \sum_{\substack{h \bmod 2k \\ \text{g.c.d.}(h,2k)=1}} \delta_1^{(f)}(\nu; h; h''; 2k) \exp(-\pi i h n/k)$$

$$\cdot \sum_{l=k_2+k}^{N+k-1} \int_{k^{-1}(l+1)^{-1}}^{k^{-1}l^{-1}} z^{\frac{1}{2}} J_1^{(f)}(\nu; 2k; 4z)\, d\phi$$

$$= \tfrac{1}{2}\sum_{k=1}^{N} k^{-1} \sum_{\nu=1}^{k} \sum_{l=N+1}^{N+k-1} \int_{k^{-1}(l+1)^{-1}}^{k^{-1}l^{-1}} z^{\frac{1}{2}} J_1^{(f)}(\nu; 2k; 4z)\, d\phi$$

$$\cdot \sum_{\substack{h \bmod 2k \\ \text{g.c.d.}(h,2k)=1 \\ N < k_2+k \leqq l}} \delta_1^{(f)}(\nu; h; h''; 2k) \exp(-\pi i h n/k).$$

The inner sum is now a truncated generalized Kloosterman sum by (4.5) and subject to the estimate of Lemma 4.1. Hence

$$|\Sigma_{23}| = O\Big(\sum_{k=1}^{N} k^{-1}\sum_{\nu=1}^{k} N^{\frac{1}{2}}k\,|\,k-2(\nu-\tfrac{1}{6})|^{-1}(kN)^{-1}(n,k)^{\frac{1}{2}}k^{\frac{1}{2}+\epsilon'}\Big)$$

$$= O\Big(N^{-\frac{1}{2}}\sum_{k=1}^{N}(n,k)^{\frac{1}{2}}k^{-\frac{1}{2}+\epsilon''}\log k\Big)$$

$$= O(N^{\epsilon'''}) \quad \text{(just as for } \Sigma_{21}\text{)}$$

$$= O(n^{\epsilon}).$$

Thus we have $|\Sigma_2| = O(n^{\epsilon})$ for any $\epsilon > 0$. Σ_3 is handled in exactly the same manner. Hence

$$A(n) = \sum_{k=1}^{[n^{\frac{1}{2}}]}\frac{\lambda(k)\exp\{\pi(n-1/24)^{\frac{1}{2}}/k6^{\frac{1}{2}}\}}{k^{\frac{1}{2}}(n-1/24)^{\frac{1}{4}}} + O(n^{\epsilon}).$$

Coupling this with (4.6.6), we obtain the required result.

In an exactly analogous manner one may obtain the following

THEOREM 5.1. *Let $M(q)$ be one of the functions ϕ, ψ, ν, ω. Then if*

$$M(q) = \sum_{n=0}^{\infty} A^{(M)}(n)q^{n},$$

$$A^{(\phi)}(n) = \sum_{\substack{k=1\\ 2\,|\,k}}^{[n^{\frac{1}{2}}]}\frac{\lambda^{(\phi)}(k)\exp\{\pi(n-1/24)^{\frac{1}{2}}/k6^{\frac{1}{2}}\}}{k^{\frac{1}{2}}(n-1/24)^{\frac{1}{4}}} + O(n^{\epsilon}),$$

$$\lambda^{(\phi)}(k) = \tfrac{1}{2}(-1)^{n+\frac{1}{2}k}A_{2k}(n-\tfrac{1}{2}k) \qquad \text{(by (4.6.7))}$$

$$A^{(\psi)}(n) = \sum_{\substack{k=1\\ (k,2)=1}}^{[n^{\frac{1}{2}}]}\frac{\lambda^{(\psi)}(k)\exp\{\pi(n-1/24)^{\frac{1}{2}}/k6^{\frac{1}{2}}\}}{2k^{\frac{1}{2}}(n-1/24)^{\frac{1}{4}}} + O(n^{\epsilon}),$$

$$\lambda^{(\psi)}(k) = \tfrac{1}{2}(-1)^{n+\frac{1}{2}(k-1)}A_{2k}(n) \qquad \text{(by (4.6.8))}$$

$$A^{(\nu)}(n) = \sum_{\substack{k=1\\ (k,2)=1}}^{[n^{\frac{1}{2}}]}\frac{\lambda^{(\nu)}(k)\exp\{\pi(n-1/3)^{\frac{1}{2}}/k6^{\frac{1}{2}}\}}{(2k)^{\frac{1}{2}}(n+1/3)^{\frac{1}{4}}} + O(n^{\epsilon}),$$

$$\lambda^{(\nu)}(k) = \tfrac{1}{2}(-1)^{n+\frac{1}{2}(k-1)}A_{k}\{(n-1)[\tfrac{1}{2}(k+1)]^{2}-3(k^{2}-1)/8\} \qquad \text{(by 4.6.9))}$$

$$A^{(\omega)}(n) = \sum_{\substack{k=1\\(k,2)=1}}^{[n^{\frac{1}{2}}]} \frac{\lambda^{(\omega)}(k)\exp\{\pi(n+2/3)^{\frac{1}{2}}/k3^{\frac{1}{2}}\}}{k^{\frac{1}{2}}(n+2/3)^{\frac{1}{2}}} + O(n^{\epsilon}),$$

$$\lambda^{(\omega)}(k) = \tfrac{1}{2}(-1)^{\frac{1}{2}(k-1)}A_k\{n\tfrac{1}{2}(k+1) - 3(k^2-1)/8\} \quad (\text{by } 4.6.10)).$$

THE PENNSYLVANIA STATE UNIVERSITY,
UNIVERSITY PARK, PENNSYLVANIA.

REFERENCES.

[1] R. Bellman, *A brief introduction to theta functions*, Holt, Rinehart and Winston, New York, 1961.

[2] L. A. Dragonette, "Some asymptotic formulae for the mock theta series of Ramanujan," *Transactions of the American Mathematical Society*, vol. 72 (1952), pp. 474-500.

[3] T. Estermann, "Vereinfachter Beweis eines Satzes von Kloosterman," *Abhandlungen Hamburg Mathematisches Seminar*, vol. 7 (1929), pp. 82-98.

[4] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 4th ed., Oxford, Oxford University Press, 1960, 421 pp.

[5] K. G. J. Jacobi, *Fundamenta nova theoriae functionum ellipticarum*, Köningsberg, 1829.

[6] D. H. Lehmer, "The series for the partition function," *Transactions of the American Mathematical Society*, vol. 43 (1938), pp. 271-295.

[7] J. Lehner, "A partition function connected with the modulus five," *Duke Journal of Mathematics*, vol. 8 (1941), pp. 631-655.

[8] H. Rademacher, "The Fourier coefficients of the modular invariant $J(\tau)$," *American Journal of Mathematics*, vol. 60 (1938), 501-512.

[9] S. Ramanujan, *Collected papers of Srinivasa Ramanujan*, Cambridge, Cambridge University Press, 1927, 355 pp.

[10] H. Salié, "Uber die Kloostermanschen Summen $S(u, v; q)$ (" *Mathematik Zeitschrift*, vol. 34 (1931), pp. 91-109.

[11] J. Tannery and J. Molk, *Elements de la Theorie des Fonctions Elliptiques*, Paris, 1896, vol. 2, 299 pp.

[12] G. N. Watson, "The final problem: an account of the mock theta functions," *Journal of the London Mathematical Society*, vol. 11 (1936), pp. 55-80.

[13] M. I. Knopp and J. R. Smart, "On Kloosterman sums connected with modular forms of half-integral dimension," *Illinois Journal of Mathematics*, vol. 8 (1964), pp. 480-487.

STRUCTURALLY STABLE SYSTEMS ARE NOT DENSE.

By S. Smale.

In this note, we give a negative answer to "the problem of structural stability"; are the structurally stable differential equations dense in the C' topology in all (first order, ordinary, autonomous) differential equations?

Main Theorem. *There exists a compact 4 dimensional manifold M, an open set U in the space of C^r vector fields, C^r topology, $r > 0$, on M such that no $X \in U$ is structurally stable.*

This problem has been stated explicitly in the above form by the author on several occasions, see e.g. [10] and [11]. However the problem is older, considered for example by Soviet mathematicians, and by Peixoto who gave an affirmative answer [7] for the 2-disk and [8] for compact 2-manifolds.

Evidence for a positive answer in higher dimensions was given by the author [12], [13] where it was shown that certain examples of differential equations having an infinite number of periodic solutions were structurally stable, and Anosov [2] gave further examples with the same properties.

The notion of structural stability was introduced by Andronov and Pontriagin in 1937 [1] (see also Lefschetz [5]). We recall the definitions now.

A *differential equation* (1st order, ordinary) X, will be a C^r tangent vector field on a C^∞ manifold M, which we will assume compact for purposes of this paper. Of course X, through its solution curves, generates a 1 parameter group of diffeomorphisms ϕ_t of M. An *equivalence* between differential equations X and X' on M is a homeomorphism $h: M \to M$ which sends a sensed solution curve of X onto one of X'. Fixing a metric on M, it is an ϵ-equivalence if it is pointwise within ϵ of the identity. The C^r differential equations on M form a normed space with the C' norm. Then X is *structurally stable* if given $\epsilon > 0$, there exists $\delta > 0$ such that if $\| X - X' \|_{C'} < \delta$, then X' is ϵ-equivalent to X.

We will first construct a diffeomorphism g of a 3-dimensional manifold (a 3-torus) and use this to construct a differential equation X_0 on a 4-dimensional manifold. This X_0 has the property that there is a neighborhood N of X_0 in the C' topology such that no X in N is structurally stable.

Received June 1, 1965.

The starting point is the linear transformation of R^3 given by the matrix

$$A = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & r \end{pmatrix}$$

where $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has determinant 1, a, b, c, d are integers and $0 < r < 1$, r small. We assume further that the eigenvalues of B are real and not ± 1. Then B as a transformation of the x-y plane induces a diffeomorphism B_0 of the 2-torus T.

Also A as a transformation on R^3 induces a diffeomorphism $A_0 \colon T \times R \to T \times R$ with $T \times 0$ invariant and all other points of $T \times R$ contracting towards $T \times 0$ under A_0.

Of course A_0 restricted to $T \times 0$ is essentially the same as B_0 on T. We shall now recall some of the structure of this discrete flow (see [2], [3], [10], R. Thom [unpublished]).

Let $\Pi \colon R^2 \to T$ be the covering map, and $p = \Pi(0,0)$ so that p is a fixed point of B_0. Then the two 1-dimensional eigenspaces of B project under Π into a stable and unstable manifold through p (see [11] for these definitions of these terms) which we will denote respectively by $W^s(p)$ and $W^u(p)$. Both $W^s(p)$ and $W^u(p)$ are dense in T.

The periodic points of B_0 are dense in T (this fact may be proved algebraically, one may use the Generalized Birkhoff Theorem of [13] or Anosov [2]), and each of these also has a 1-dimensional stable (and unstable) manifold associated to it. These will be disjoint, parallel (using the Euclidean structure on R^2) to $W^s(p)$ and each is dense in T.

Now we pass to the extension $A_0 \colon T \times R \to T \times R$ of B_0. It is easy to see that the periodic points of A_0 coincide with those of B_0 as do also the 1-dimensional unstable manifolds which we will denote from now on by $W_1{}^u(q_i)$, $i = 1, 2, \cdots, q_i$ the periodic points of A_0.

The stable manifolds now however are two dimensional, being in $T \times R$ of the form $W^s(q_i) \times R$ where $W^s(q_i) \subset T$ is the 1-dimensional stable manifold of $B_0 \colon T \to T$ for the periodic point q_i. From now on we denote this 2-dimensional stable manifold $W^s(q_i) \times R$ by $W_2{}^s(q_i)$.

Thus the stable manifolds of the periodic points of A_0 are a countable, dense family of parallel "planes" in $T \times R$.

Let $\bar{A}_0$ denote the restriction of A_0 to $T \times [-1, 1] \subset T \times R$. In a neighborhood S of radius $\frac{1}{2}$ about the point $b = (0, 0, 2)$ of R^3, let C be a linear transformation which contracts in the x-y plane through b, leaving b

fixed and leaving the z-axis invariant, expanding on it away from b. Denote by C_0 the induced transformation defined on the corresponding neighborhood $S_0 = \Pi_1(S)$ of $\Pi_1(b)$ in $T \times R$, $\Pi_1 \colon R^3 \to T \times R$ our covering map, and let $J = \{(0,0,z) \in T \times R \mid 0 \leqq z \leqq 2\}$.

Now it is easy to define a diffeomorphism f on a neighborhood of $T \times [-1,1] \cup S_0 \cup J$ in $T \times R$ into $T \times R$ which is $\bar{A}_0$ on $T \times [-1,1]$, C_0 on S_0, and leaves J invariant with no fixed points on J, except 0 and $\Pi_1(b)$.

In fact we may assume f is extended not just to a neighborhood of $T \times [-1,1] \cup S_0 \cup J$, but to a diffeomorphism of the whole compact manifold $T \times S' \xrightarrow{f} T \times S'$, where S' is $[-3,3]$ with end points identified.

Our desired diffeomorphism g is obtained by perturbing f in a neighborhood G of a point α in $T \times S'$ as follows. We assume $r = \frac{1}{2}$. Let $\alpha_0 = \Pi_1(0,0,\frac{3}{4})$ and $G_0 = \Pi_1(N_{1/8}(0,0,\frac{3}{4}))$, $N_\epsilon(x)$ being the neighborhood of x of radius ϵ. Then let $\alpha = f^{-1}(\alpha_0)$ and $G = f^{-1}(G_0)$ so since $r = \frac{1}{2}$, $f(G) \cap G = \phi$.

Using the obvious linear structure on G, let $g = f$ outside G and on G let $g = f + \eta\phi h$ where h is the vector $(1,0,0)$, ϕ is a non-negative C^∞ function with compact support in G and non-degenerate maximum value 1 at α and finally $\eta > 0$ is small enough so that g is a diffeomorphism. Note h is transversal to the "planes" $W_2{}^s(q_i)$.

We now show that if $g' \colon M \to M$ is any diffeomorphism sufficiently close to g in the C' topology, then there exists a diffeomorphism $g'' \colon M \to M$ arbitrarily C' close to g' with the property that g' and g'' are not topologically conjugate, at least by a homeomorphism pointwise close to the identity. (Thus g' is not a *structurally stable diffeomorphism*; see [10], [11] for elaboration of this).

One can best study a small C' perturbation g' of g by using theorems of Perron [9] put into a general setting in the spirit of Anosov [2]. We expect to expand on this point in a future paper. In the meantime we argue as follows. First, for g' C^1-close enough to g, there is an invariant torus $T' \subset M = T \times S'$ corresponding to $T \times 0$ and C'-close to it. For this see Kupka [4] or Moser [6]. Next, g' restricted to T' is topologically conjugate by a homeomorphism close to the identity to $g \colon T \times 0 \to T \times 0$; see [2] or [3]. Let the stable manifolds of g' of the countable dense set of periodic points q_i' on T' be denoted by $W_2{}^s(q_i')'$. The tangent 2-plane to each point of $W_2{}^s(q_i')'$ is close (arbitrarily, depending on the C' distance of g' to g) to the plane $W_2{}^s(q_i)$ (which is independent of i, of course) on $T \times [-1,1]$.

This is a consequence of Section 5 of [13]. Similarly, the $W_2{}^s(q_i')'$ filled up $T \times [-1,1]$ densely.

The intersections of $W_2{}^s(q_i')'$ near the z-axis with the x-axis will be a countable dense set Σ'. Denote by $W_1{}^u$ the 1-dimensional unstable manifold of the fixed point $(0,0,2)$ of g and by $W_1{}^{u'}$ that corresponding to the perturbation g'.

Now for all g' (and g'') sufficiently close to g, we can have precisely one of two cases as follows.

Case 1: The first point on $W_1{}^{u'}$, near the bump $f(\alpha)$, ordering by Σ' is on some $W_2{}^s(q_i')'$.

Case 2: It is not.

The following two facts finish our picture of g'.

1. By letting $g'' = g' + \eta\phi h$ (see the definition of g), arbitrarily small η, we can suppose g'' is in Case 2 if g' is of Case 1 and vice versa.

2. If g' and g'' are in the opposite cases there is no homeomorphism h of M close to the identity such that $g'h = hg''$. This can be seen by iterating g' and g''.

From any diffeomorphism g of a compact n-manifold M onto itself, one can construct in a canonical way, an ordinary differential equation X on an $(n+1)$ manifold with a cross-section $g\colon M \to M$ (see [11]). In this way the various relevant properties of g correspond to those of X. Therefore, from what we have shown, the main theorem follows.

The example in this paper surely reduces the importance of the notion of structural stability. One might be further discouraged from studying the global qualitative theory of higher dimensional, ordinary differential equations. We believe, however, that this study can be constructive and are preparing a paper in this direction based on an axiom, which we consider to be of central importance, axiom A, described as follows.

Let X be a smooth vector field on a compact manifold M which generates a 1-parameter flow $\phi_t\colon M \to M$. Relative to ϕ_t a point $p \in M$ is called *wandering* if there is a neighborhood U of p such that $\phi_t(U) \cap U = \emptyset$ for all t with $|t|$ greater than some t_0. We will denote the non-wandering points by $\Omega_X = \Omega \subset M$. Then Ω is a closed invariant subset of M.

If $\Pi\colon E \to M$ is a vector space bundle over a topological space, a 1-*parameter group of V-B maps* is a 1-parameter group of homeomorphisms $\psi_t\colon E \to E$ such that for each $t \in R$, ψ_t is a V-B map (fiber preserving and linear on each fiber). For example if $\phi_t\colon M \to M$ is a smooth 1-parameter flow on a differentiable manifold, the derivative $\tilde{\phi}_t\colon T_M \to T_M$ defines a 1-

parameter group of the *V-B* maps of the tangent bundle, called the *derived flow.*

A 1-parameter group of *V-B* maps $\psi_t\colon E \to E$ over a compact space M is called *contracting* if for some Riemannian structure on $\Pi\colon E \to M$, the following two estimates hold:

$$1) \quad \| \psi_t(\nu) \| \leqq a \| \nu \| e^{-ct} \quad , \quad t \geqq 0, \nu \in E,$$

$$2) \quad \| \psi_t(\nu) \| \geqq b \| \nu \| e^{-ct} \qquad t \leqq 0, \nu \in E.$$

Here a, b, c are positive constants, and $\| \nu \|$ is the norm of ν in the metric on the fiber to which ν belongs.

Finally ψ_t is called *expanding* if ψ_{-t} is contracting.

For any closed invariant set $\Omega \subset M$ of a smooth differentiable flow $\phi_t\colon M \to M$, a *U-structure* on Ω (generalizing Anosov [2]) is defined as follows. The tangent bundle $T(M)$ of M restricted to Ω, $T_\Omega(M)$, is a Whitney sum $T_\Omega(M) = D_\Omega{}^u \oplus D_\Omega{}^s \oplus \dot{X}$ (topologically) such that

$$E_\Omega{}^u = D_\Omega{}^u + \dot{X}, \qquad E_\Omega{}^s = D_\Omega{}^s + \dot{X},$$

are invariant under the derived flow $\tilde{\phi}_t\colon T(M) \to T(M)$ of ϕ_t, $\tilde{\phi}_t\colon E_\Omega{}^s \to E_\Omega{}^s$ is contracting (as a 1-parameter group of *V-B* maps) and $\tilde{\phi}_t\colon E_\Omega{}^u \to E_\Omega{}^u$ is expanding. Here $\dot{X}$ is the line bundle defined by X, X the tangent vector field generating ϕ_t.

For a smooth flow $\phi_t\colon M \to M$ of a compact manifold then our axiom is:

Axiom A. The set of non-wandering points $\Omega \subset M$ of ϕ_t has a *U*-structure.

REFERENCES.

[1] A. A. Andronov and L. S. Pontriagin, "Systimes Grossiers," *Doklady Akademii Nauk*, vol. 14 (1937), pp. 247-251.

[2] D. V. Anosov, "Roughness of geodesic flows on compact Riemannian manifolds of negative curvature," *Soviet Mathematics*, vol. 3 (1962), pp. 1068-1070.

[3] V. I. Arnold and Ja. G. Sinai, "Small perturbations of the automorphisms of the torus," *Soviet Mathematics*, vol. 3 (1962), pp. 783-787. See also correction of same, *Soviet Mathematics*, vol. 4 (1963) page preceding 561.

[4] I. Kupka, "Stabilite des varietes invariantes d'un champ de vecteurs pour les petites perturbations," *Comptes Rendus de l'Academie des Sciences de Paris*, vol. 258 (1964), pp. 4197-4200.

[5] S. Lefschetz, *Differential Equations: Geometric Theory*, New York, 1957.

[6] J. Moser, *On invariant manifolds of vector fields and symmetric partial differential equations*, in Differential Analysis, Bombay, 1964.

[7] M. Peixoto, "On structural stability," *Annals of Mathematics*, vol. 69 (1959), pp. 199-122.

[8] ———, "Structural stability on 2-dimensional manifolds," *Topology*, vol. 2 (1962), pp. 101-121.

[9] Perron, "Die Stabilitatsfrage bei Differentialgleichungen," *Mathematische Zeitschrift*, vol. 32 (1930), pp. 703-728.

[10] S. Smale, *Dynamical systems and the topological conjugacy problem for diffeomorphisms*, Proceedings of the International Congress of Mathematicians, Stockholm, 1962.

[11] ———, "Stable manifolds for differential equations and diffeomorphisms," *Annali della Scuola Normale Superiore di Pisa*, Series III, vol. XVII (1963), pp. 97-116.

[12] ———, *A structurally stable differentiable homeomorphism with an infinite number of periodic points*, Report on the Symposium on non-linear oscillations, Kiev Mathematical Institute, 1961.

[13] ———, *Diffeomorphisms with many periodic points*, M. Morse Symposium, Differential and Combinatorial Topology, Princeton, 1965.

THE MODULAR GROUPS OF HILBERT AND SIEGEL.*

By William F. Hammond.

Introduction. It has only been within the last five years that one has begun to understand completely even the simplest cases of modular functions of several variables. In connection with the modular functions of Siegel, J. Igusa [7] has determined the structure of the graded ring of modular forms in the genus two case. Around the same time, K.-B. Gundlach [3] settled the same question for the Hilbert modular functions arising from the quadratic algebraic number field of discriminant 5 by an entirely different method. Gundlach's method is a generalization of a classical technique for studying elliptic modular forms, which has been shown recently [4] to work also in the genus two Siegel case. The present investigations arose from some suggestions made to the author by Igusa, who knew that in certain special cases the Hilbert modular group of a real quadratic field could be imbedded in the Siegel modular group of genus two. The idea was, then, to investigate such imbeddings with an eye toward using known results concerning Siegel modular forms of genus two for the study of Hilbert modular forms.

We shall proceed to outline the contents of this paper. In § 1 we recall a few basic definitions. In § 2 the notion of "modular imbedding" for a totally real algebraic number field is introduced. Roughly, for a field of degree n, a modular imbedding gives an analytic mapping of the domain on which Hilbert's modular group operates into the domain on which the Siegel modular group of genus n operates in such a way that the action of the Hilbert modular group is carried into the action of the Siegel modular group and functions belonging to the Siegel modular group are "pulled back" into functions belonging to the Hilbert modular group. It is shown that modular imbeddings exist for a given totally real algebraic number field if and only if the narrow ideal class of the different is a square, and the number of equivalence classes of modular imbeddings is determined. Our original investigations on modular imbeddings were confined to the quadratic case, and the

---

Received May 18, 1965.

* This research is the basis of the author's doctoral dissertation submitted to The Johns Hopkins University in 1965. The author was supported in part as a research assistant under a grant from the National Science Foundation.

above results are due to Igusa, who has kindly consented to the publication of his results here. § 3 deals with modular imbeddings for quadratic fields. It is shown, in particular, that modular imbeddings exist for a given quadratic field if and only if the discriminant of the field is the sum of two squares.

The results of §§ 4, 5, 6 have meaning only for quadratic fields which admit modular imbeddings. In § 4 we obtain without effort the existence of algebraically independent Hilbert modular forms of weights two, six, and ten. As far as we know, no general statement of this type has been known. If a modular imbedding for a quadratic field is given, then (§ 5) the "diagonal" subvariety of the Siegel modular variety of genus two induces a "bunch of curves" on the Hilbert modular variety of the quadratic field. The said bunch of curves is irreducible if and only if the field is that of discriminant 5. These considerations yield a new proof of Gundlach's result on the structure of the graded ring of symmetric Hilbert modular forms of even weight for the field of discriminant 5, and they enable us to obtain (§ 6) a similar result for the field of discriminant 8.

1. Preliminaries.

Let A be a domain of integrity. The *symplectic group of degree g over A* is the group $Sp(g,A)$ consisting of all $2g$-by-$2g$ matrices M with coefficients in A satisfying the condition ${}^tMIM = I$ where ${}^tM$ denotes the transpose of M and I stands for the matrix

$$I = \begin{pmatrix} -1_g & 0 \\ 0 & 1_g \end{pmatrix}$$

with 1_g denoting the g-by-g identity matrix. The group $Sp(g,\boldsymbol{R})$, where $\boldsymbol{R}$ is the field of real numbers, operates in a natural manner on a domain $\mathfrak{S}_g$, the *Siegel upper-half plane of genus g*, which is defined as the set of all symmetric complex g-by-g matrices with positive definite imaginary parts. If τ is in $\mathfrak{S}_g$ and if M in $Sp(g,R)$ has the form

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

then the image $M\cdot\tau$ of τ under M is given by $M\cdot\tau = (a\tau + b)(c\tau + d)^{-1}$. The group $Sp(g,\boldsymbol{Z})$, $\boldsymbol{Z}$ denoting the ring of integers, is a discrete subgroup of $Sp(g,\boldsymbol{R})$ which, therefore, operates discontinuously on $\mathfrak{S}_g$. The group of mappings corresponding to $Sp(g,\boldsymbol{Z})$ is called the *Siegel modular group of genus g*.

Definition 1.1. A *Siegel modular form of weight w* is a holomorphic function in $\mathfrak{S}_g$ which for every M in $Sp(g,\boldsymbol{Z})$ satisfies

$$f(M\cdot\tau)=\det(c\tau+d)^w f(\tau)$$

identically for τ in $\mathfrak{S}_g$. If $g=1$ it is also required that $f(iy)$ approaches a finite limit as y approaches infinity through positive real values.

Let K be a totally real algebraic number field of degree $n>1$. It is common to imagine K as a subset of real Euclidean n-space $\boldsymbol{R}^n$ by assigning to each element ρ of K the n-tuple of its conjugates $(\rho_1,\rho_2,\cdots,\rho_n)$. In this connection it is convenient to regard complex n-space $\boldsymbol{C}^n$ as a ring with addition and multiplication defined component-wise. The norm $Nm(z)$ of a point z in $\boldsymbol{C}^n$ is the product of its coordinates. By assigning to a matrix μ in $Sp(1,K)$ the n-tuple $(\mu_1,\mu_2,\cdots,\mu_n)$ of its conjugate matrices μ_i, we obtain $Sp(1,K)$ as a subgroup of the n-fold product group $Sp(1,\boldsymbol{R})^n$. If we let $Sp(1,\boldsymbol{R})^n$ operate component-wise on the n-fold product domain $(\mathfrak{S}_1)^n$, then we obtain an operation of $Sp(1,K)$ on $(\mathfrak{S}_1)^n$. Since the ring $\boldsymbol{\mathcal{O}}$ of integers in K is a discrete subset of $\boldsymbol{R}^n$, we see that the group $Sp(1,\boldsymbol{\mathcal{O}})$ operates discontinuously on $(\mathfrak{S}_1)^n$, and the corresponding group of mappings is called the *Hilbert modular group of K.*

Definition 1.2. A *Hilbert modular form of weight w for K* is a holomorphic function in $(\mathfrak{S}_1)^n$ which for every μ in $Sp(1,\boldsymbol{\mathcal{O}})$ satisfies

$$f(\mu\cdot\tau)=Nm(\gamma\tau+\delta)^w f(\tau)$$

identically for τ in $(\mathfrak{S}_1)^n$, where $(\gamma\,\delta)$ is the bottom row of μ. A Hilbert modular form will be called *symmetric* if its value at a given point of $(\mathfrak{S}_1)^n$ is not changed by permutation of its coordinates.

2. Modular imbeddings. If z is an element of $\boldsymbol{C}^n$, we shall use z^* to denote the n-by-n diagonal matrix whose diagonal coefficients are the coordinates of z. If $m=(m_1,\cdots,m_n)$ is an element of $Sp(1,\boldsymbol{R})^n$ with

$$m_i=\begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix},$$

then we put $a=(a_1,\cdots,a_n)$ and likewise for b, c, and d; by m^* we shall mean the element of $Sp(n,\boldsymbol{R})$ given by

$$m^*=\begin{pmatrix} a^* & b^* \\ c^* & d^* \end{pmatrix}.$$

Let ϕ_0 denote the holomorphic imbedding of $(\mathfrak{S}_1)^n$ in $\mathfrak{S}_n$ defined by mapping τ to τ^*, and let Φ_0 denote the monomorphism of $Sp(1,\boldsymbol{R})^n$ in $Sp(n,\boldsymbol{R})$ defined by mapping m to m^*. Then obviously $\phi_0(m\cdot\tau)=\Phi_0(m)\cdot\phi_0(\tau)$. If N is an arbitrary element of $Sp(n,\boldsymbol{R})$, and if we let $\phi(\tau)=N\cdot\phi_0(\tau)$ and $\Phi(m)$

$= N\Phi_0(m)N^{-1}$, then we also have $\phi(m\cdot\tau) = \Phi(m)\cdot\phi(\tau)$. We shall say that (ϕ, Φ) defined in this way is an *alteration* of (ϕ_0, Φ_0). Let K be a totally real algebraic number field of degree n.

Definition 2.1. A *modular imbedding for* K is a pair (ϕ, Φ) consisting of a holomorphic injection ϕ of $(\mathfrak{S}_1)^n$ in $\mathfrak{S}_n$ together with a monomorphism Φ of $Sp(1, \boldsymbol{R})^n$ in $Sp(n, \boldsymbol{R})$ such that:

(i) (ϕ, Φ) is an alteration of (ϕ_0, Φ_0).

(ii) The image under Φ of $Sp(1, \mathcal{O})$ is contained in $Sp(n, \boldsymbol{Z})$.

(iii) If f is a Siegel modular form of weight w, then the composition $f\phi$ is a Hilbert modular form of weight w for K.

For M in $Sp(n, \boldsymbol{R})$ with n-by-n components a, b, c, d we define $R(M, \tau)$ for τ in $\mathfrak{S}_n$ by $R(M, \tau) = \det(c\tau + d)$. From the fact that f is a Siegel modular form of weight w we see that $f\phi$ satisfies a functional equation under the operations of $Sp(1, \mathcal{O})$ with multiplier $R(\Phi\mu, \phi\tau)^w$. Therefore, the condition (iii) above can be replaced[1] by the condition

(iv) $R(\Phi\mu, \phi\tau) = R(\mu^*, \tau^*)$.

PROPOSITION 2.2. *If the real symplectic matrix N with n-by-n components a, b, c, d defines (ϕ, Φ) as an alteration of (ϕ_0, Φ_0) so that (ϕ, Φ) satisfies the condition* (ii), *then a necessary and sufficient condition that (ϕ, Φ) be a modular imbedding for K is that the left bottom n-by-n component c of N be zero.*

Proof. Since we are assuming the condition (ii), we show that the condition (iv) is satisfied if and only if $c = 0$. Now (iv) is satisfied if and only if for each μ in $Sp(1, \mathcal{O})$ we have $R(N, (\mu\cdot\tau)^*) = R(N, \tau^*)$. This is certainly the case if $c = 0$; conversely we suppose that this holds. This means that the function h on $(\mathfrak{S}_1)^n$ defined by $h(\tau) = R(N, \tau^*)$ is invariant under all of the Hilbert modular mappings. It is well known that h must then be constant since the only holomorphic modular invariants are the constants. We have $h(\tau) = h(0) = \det(d) \neq 0$ since $h(\tau)$ is non-vanishing in $(\mathfrak{S}_1)^n$. As a result, $\det(ts + 1_n) = 1$ for every complex scalar t, where s is the symmetric matrix $d^{-1}c$. It is sufficient to show that s is zero. If s were not zero, then it would have a non-zero eigenvalue t, and then $\det(t1_n - s) = 0$ is a contradiction.

---

[1] **Please see the addendum at the end of this paper.**

Definition 2.3. Two modular imbeddings (ϕ_1, Φ_1) and (ϕ_2, Φ_2) will be called *equivalent* if there is an element M of $Sp(n, \mathbf{Z})$ such that $\phi_2 = M\phi_1$ and $\Phi_2 = M\Phi_1 M^{-1}$.

Let N_1 and N_2 be elements of $Sp(n, \mathbf{R})$. A necessary and sufficient condition that N_1 and N_2 define the same modular imbedding (ϕ, Φ) is that there exists $p = (p_1, \cdots, p_n)$, with each p_i either 1 or -1, such that $N_2 = N_1 P$, where P is the element of $Sp(n, \mathbf{Z})$ defined by

$$P = \begin{pmatrix} p^* & 0 \\ 0 & p^* \end{pmatrix}.$$

Thus, N_1 and N_2 define equivalent modular imbeddings if and only if there is an element M of $Sp(n, \mathbf{Z})$ and a matrix P of the above type such that $N_2 = MN_1P$.

PROPOSITION 2.4. *Every modular imbedding (ϕ, Φ) is equivalent to a modular imbedding (ϕ_1, Φ_1) in which ϕ_1 is homogeneous linear.*

Proof. Let N be a real symplectic matrix which defines (ϕ, Φ) as an alteration of (ϕ_0, Φ_0). Then $N = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ can be decomposed as the product SN_1 where $S = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$ and $N_1 = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}$ with $s = b^t a = \phi(0)$. We shall show that s is necessarily an integer matrix. Once this is clear it will be obvious that the alteration (ϕ_1, Φ_1) of (ϕ_0, Φ_0) defined by N_1 is a modular imbedding which is equivalent to the original modular imbedding. Since (ϕ, Φ) is a modular imbedding, the condition (ii) of Definition 2.1 requires that the following matrices have integer coefficients:

$$\Phi\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a^t a \\ 0 & 1 \end{pmatrix}, \qquad \Phi\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} * & -b^t b \\ d^t d & * \end{pmatrix},$$

$$\Phi\begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} * & a^t a + b^t b + b^t a. \\ * & * \end{pmatrix}.$$

Hence, $s = b^t a$ has integer coefficients. This completes the proof of Proposition 2.4. We note, however, that $a^t a$ and $d^t d$ are integer matrices. Since they are inverses of each other, we see that each is a positive definite symmetric integer matrix with determinant 1. Let $\psi(z) = az^* a^{-1}$ for z in $\mathbf{C}^n$, and let $\boldsymbol{u} = a^t a$. Then ψ is a normal repesentation of K (in the sense that elements of $\mathfrak{O}$ are represented by integer matrices), and ψ and $\boldsymbol{u}$ are related by the condition ${}^t\psi = \boldsymbol{u}^{-1}\psi\boldsymbol{u}$. The normality is a consequence of condition (ii) of Definition 2.1 together with the formula

$$\Phi_1\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}=\begin{pmatrix}\psi(\alpha) & \psi(\beta)\boldsymbol{u}\\ \boldsymbol{u}^{-1}\psi(\gamma) & {}^t\psi(\delta)\end{pmatrix}. \tag{2.5}$$

PROPOSITION 2.6. *There is a one-to-one correspondence between homogeneous linear modular imbeddings* (ϕ,Φ) *for* K *and pairs* $(\psi,\boldsymbol{u})$ *consisting of a non-degenerate normal representation* ψ *of* K *by rational matrices of degree* n *and an* n*-by-*n *symmetric positive definite integer matrix* $\boldsymbol{u}$ *with determinant* 1 *such that for all elements* ρ *of* K *the relation* ${}^t\psi(\rho)=\boldsymbol{u}^{-1}\psi(\rho)\boldsymbol{u}$ *holds.*

Proof. The transformation from (ϕ,Φ) to $(\psi,\boldsymbol{u})$ has been described above, and it is obviously one-to-one "into." The point which must be settled is this: Suppose that $(\psi,\boldsymbol{u})$ is given, ψ is extended $\boldsymbol{C}$-linearly to $\boldsymbol{C}^n$, ϕ is defined by $\phi(z)=\psi(z)\boldsymbol{u}$, and Φ is defined by (2.5) for elements of $Sp(1,\boldsymbol{R})^n$; can we conclude that there is a non-singular real matrix a such that $\psi(z)=az^*a^{-1}$ and $\boldsymbol{u}=a{}^ta$? The answer is affirmative. There is a non-singular real matrix W such that $\psi(\rho)=W\rho^*W^{-1}$. The relation between ψ and $\boldsymbol{u}$ implies that the matrix $P={}^tW\boldsymbol{u}^{-1}W$ commutes with ρ^* for every ρ in K, and therefore P is a diagonal matrix. Since P is positive definite, we may write $P=P_0^2$ for some real diagonal matrix P_0. Then $a=WP_0^{-1}$ meets our requirements.

Remark 2.7. $(\psi_1,\boldsymbol{u}_1)$ and $(\psi_2,\boldsymbol{u}_2)$ correspond to equivalent homogeneous linear modular imbeddings if and only if there is an n-by-n integer unimodular matrix $\boldsymbol{v}$ such that $\psi_2=\boldsymbol{v}\psi_1\boldsymbol{v}^{-1}$ and $\boldsymbol{u}_2=\boldsymbol{v}\boldsymbol{u}_1{}^t\boldsymbol{v}$.

THEOREM 2.8 (Igusa). *A totally real algebraic number field* K *admits modular imbeddings if and only if the narrow ideal class of the different is a square in the narrow ideal class group of* K.

Proof. Let us recall (cf. [5]) that the *different* of K is the reciprocal $(\mathfrak{O}')^{-1}$ of the fractional ideal $\mathfrak{O}'$ which is complementary to the unit ideal and that two ideals (fractional) are *narrowly equivalent* if one is a totally positive multiple of the other. Suppose that $(\psi,\boldsymbol{u})$ corresponds to some homogeneous linear modular imbedding. There is a column vector ω, whose entries constitute a $\boldsymbol{Z}$-basis for some ideal $\mathfrak{A}$ of K, such that $\psi(\rho)\omega=\omega\rho$ for all ρ in K. Let W be the non-singular real matrix whose columns are the conjugate columns of ω. Then $\psi=W(\ )^*W^{-1}$, and the matrix $P={}^tW\boldsymbol{u}^{-1}W$, as in the proof of (2.6), is a positive definite diagonal matrix. It is clear that $P=\rho^*$ where ρ is the totally positive element ${}^t\omega\boldsymbol{u}^{-1}\omega$ of K. We have $u_{ij}=\mathrm{trace}(\rho^{-1}\omega_i\omega_j)$, and since the products $\omega_i\omega_j$ generate $\mathfrak{A}^2$ as an additive group, it follows that the principal ideal $\rho^{-1}\mathfrak{O}$ is contained in the ideal $(\mathfrak{A}^2)'$ complementary to $\mathfrak{A}^2$.

Taking norms, we see that $\rho^{-1}\mathfrak{O} = (\mathfrak{A}^2)'$, or $\mathfrak{O}' = \rho^{-1}\mathfrak{A}^2$, which shows that the narrow class of the different is a square. This reasoning is reversed without difficulty.

Thus, homogeneous linear modular imbeddings for K correspond (in a many-to-many fashion) with pairs $(\mathfrak{A}, \rho)$ where $\mathfrak{A}$ is a fractional ideal of K and ρ is a totally positive element of K such that $\mathfrak{A}^2 = \rho\mathfrak{O}'$. We shall say that two such pairs $(\mathfrak{A}_1, \rho_1)$ and $(\mathfrak{A}_2, \rho_2)$ are *equivalent* if and only if there is a non-zero element α of K such that $\mathfrak{A}_2 = \alpha\mathfrak{A}_1$ and $\rho_2 = \alpha^2\rho_1$. The correspondence between pairs $(\psi, \boldsymbol{u})$ and pairs $(\mathfrak{A}, \rho)$ induces a one-to-one correspondence between the equivalence classes of $(\psi, \boldsymbol{u})$ (cf. 2.7) and the equivalence classes of $(\mathfrak{A}, \rho)$; therefore, *the number of classes of modular imbeddings is equal to the number of classes of* $(\mathfrak{A}, \rho)$. We note that if $\mathfrak{A}$ is a fractional ideal of K such that $\mathfrak{A}^2$ is narrowly equivalent to a given ideal $\mathfrak{B}$, then the same thing is true for any ideal in the *usual* ideal class of $\mathfrak{A}$. Thus, one may speak of the number $\pi(K)$ of usual ideal classes whose squares are narrowly equivalent to the narrow class of $\mathfrak{O}'$. (This number is zero if there are no such classes.)

THEOREM 2.9 (Igusa). *The number of equivalence classes of modular imbeddings for a totally real algebraic number field K is the product of $\pi(K)$ with the index of the subgroup of squares of units in the group of totally positive units.*

Proof. Select inequivalent ideals $\mathfrak{A}_1, \cdots, \mathfrak{A}_\pi$ whose squares are narrowly equivalent to $\mathfrak{O}'$. The number of classes of pairs of the form $(\mathfrak{A}_k, \rho)$ for fixed k is the said index since, if $(\mathfrak{A}_k, \rho_0)$ is a fixed such pair, then $\rho = \eta\rho_0$ where η is a totally positive unit.

3. The quadratic case.

One would hope that the results of the preceding section had simple interpretations in the case $n = 2$. This is indeed the case as the following proposition shows:

PROPOSITION 3.1. *If K is a real quadratic field, then a necessary and sufficient condition that the narrow ideal class of the different of K be a square is that the discriminant D of K be the sum of two integer squares.*

Proof. If there is an ideal $\mathfrak{A}$ such that $\mathfrak{A}^2 = \rho\mathfrak{O}'$, with totally positive ρ, then $D(Nm\mathfrak{A})^2 = Nm(\rho)$, since $Nm(\rho)$ is positive, and from this we see that $1/D$, and therefore D, is the norm of a Gaussian number. It follows that D is the norm of a Gaussian integer. Conversely, if $D = u^2 + v^2$, with, say,

v even, then we choose $\mathfrak{A}$ to be the additive subgroup of K generated by v and $u + D^{\frac{1}{2}}$. Then $\mathfrak{A}$ is an ideal for which $\mathfrak{A}^2$ is narrowly equivalent to $\boldsymbol{\mathfrak{O}}'$.

Definition 3.2. Let K be an arbitrary totally real algebraic number field. A homogeneous linear modular imbedding (ϕ, Φ) for K is said to be *orthogonal* if $\phi(1) = 1_n$.

In the first place, this means that if N is any element of $Sp(n, \boldsymbol{R})$ which defines (ϕ, Φ), then $N = \begin{pmatrix} e & 0 \\ 0 & e \end{pmatrix}$ for some orthogonal matrix e. Secondly, the homogeneous linear modular imbedding corresponding to a pair $(\psi, \boldsymbol{u})$ is orthogonal if and only if $\boldsymbol{u} = 1_n$. Therefore, there is a one-to-one correspondence between orthogonal modular imbeddings for K and non-degenerate normal representations of K by *symmetric* rational matrices of degree n.

THEOREM 3.3. *If K is a totally real algebraic number field of degree n not greater than 7, then every modular imbedding for K is equivalent to an orthogonal modular imbedding.*

Proof. Let (ϕ, Φ) be a given modular imbedding for K. By (2.4) we may assume that it is homogeneous linear, and therefore it corresponds to a pair $(\psi, \boldsymbol{u})$ in the sense of (2.6). Since n is not greater than 7, there is an integer unimodular matrix $\boldsymbol{v}$ such that $\boldsymbol{v}\boldsymbol{u}{}^t\boldsymbol{v} = 1_n$. This result is essentially due to Hermite—see Dickson's *History*, [2], p. 235. If we replace ψ by $\psi_1 = \boldsymbol{v}\psi\boldsymbol{v}^{-1}$, then we obtain an orthogonal modular imbedding which is equivalent to (ϕ, Φ).

THEOREM 3.4. *Let K be the real quadratic field of discriminant D. The orthogonal modular imbeddings for K correspond in a one-to-one manner with ordered pairs (u, v) of integers such that:*

1) $D = u^2 + v^2$.

2) *v is even.*

Proof. Since the orthogonal modular imbeddings for K correspond biuniquely with the non-degenerate normal representations of K by rational symmetric 2-by-2 matrices, it suffices to establish a correspondence between these representations and the pairs (u, v). If ψ denotes such a representation, then the formula

$$\psi(D^{\frac{1}{2}}) = \begin{pmatrix} u & v \\ v & -u \end{pmatrix} \tag{3.5}$$

determines a pair (u, v), as one easily checks by observing that $\operatorname{trace}\psi(D^{\frac{1}{2}}) = \operatorname{trace}(D^{\frac{1}{2}}) = 0$ and $\det\psi(D^{\frac{1}{2}}) = \operatorname{norm}(D^{\frac{1}{2}}) = -u^2 - v^2 = -D$. Since the representative of the element $\frac{1}{2}(D + D^{\frac{1}{2}})$ of $\mathfrak{O}$ must be an integer matrix, the integer v is necessarily even. Conversely, if a pair (u, v) of integers is given which satisfies the conditions (1) and (2), then there is a unique symmetric normal representation ψ for which the formula (3.5) holds. Let us note that in any representation of D as the sum of two squares, since D is congruent to 0 or $1 \bmod 4$, at least one of the squares must be even.

It is not difficult to see that the modular imbedding corresponding to a pair of integers (u, v) is equivalent to the modular imbedding corresponding to (u', v') if and only if $(u')^2 = u^2$ and $(v')^2 = v^2$. Thus, in the quadratic case, in view of (3.3) and (3.4), the number of classes of modular imbeddings for K can be determined by counting representations of D as the sum of two squares. This is carried out, for example, in the number theory text of Hardy and Wright. On the other hand, using results from Hecke's *Vorlesungen*, [5], one can compute the number $\pi(K)$ which appears in (2.9). If we observe that D, the discriminant of K, is representable as the sum of two squares if and only if it is divisible by no prime of the form $4m + 3$, then we obtain:

THEOREM 3.6. *Let K be the real quadratic field of discriminant D, and let t be the number of primes which divide D. Modular imbeddings exist for K if and only if D is divisible by no prime of the form $4m + 3$, and in this case the number of equivalence classes of modular imbeddings for K is 2^{t-1}.*

In the case of orthogonal modular imbeddings (ϕ, Φ) we have $\phi = \psi$, and from (2.5) we see that Φ is given by

$$\Phi\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} \phi(\alpha) & \phi(\beta) \\ \phi(\gamma) & \phi(\delta) \end{pmatrix}.$$

Now let the orthogonal modular imbedding corresponding to some pair (u, v) for the real quadratic field K of discriminant D be fixed. Since $(1, 1)$ and $D^{\frac{1}{2}}(1, -1)$ are linearly independent, their images under ϕ span the image under ϕ of $\boldsymbol{C}^2$. Thus the image of $\boldsymbol{C}^2$ is the set of all symmetric complex matrices $z = \begin{pmatrix} z_1 & z_{12} \\ z_{12} & z_2 \end{pmatrix}$ such that

$$z_{12} = \tfrac{1}{2}(v/u)(z_1 - z_2), \tag{3.7}$$

and the image of K is the set of rational points on this plane. Since ϕ

preserves trace and norm, we see that the image of $\mathcal{O}$ is the set of integral points on the plane (3.7). Finally, the image under Φ of $Sp(1,\mathcal{O})$ in $Sp(2,\mathbf{Z})$ consists of all elements $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ of $Sp(2,\mathbf{Z})$ for which a, b, c, d are symmetric matrices lying in the plane (3.7).

With (u,v) still fixed, we consider the possibility of extending the map Φ to the permutation automorphism σ of $(\mathfrak{S}_1)^2$, which is defined by $\sigma(\tau_1,\tau_2)=(\tau_2,\tau_1)$. What we require is an element M of $Sp(2,\mathbf{Z})$ such that $\phi(\sigma(\tau))=M\cdot\phi(\tau)$. Since ϕ maps the fixed points of σ to the scalar points of $\mathfrak{S}_2$, the scalar points of $\mathfrak{S}_2$ must be fixed points of M. Hence $M=\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ where a is one of the eight 2-by-2 orthogonal integer matrices. Since $\phi(\sigma(z))a=a\phi(z)$ identically in z, we may let $z=D^{\frac{1}{2}}(1,-1)$ and then obtain $a=\pm j$ where $j=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. As a result, if f is a Siegel modular form of genus two, then $f\phi$ is a *symmetric* Hilbert modular form for K.

4. Hilbert modular forms. One reason for interest in modular imbeddings is that they enable us to construct Hilbert modular forms from Siegel modular forms. We shall consider only the quadratic case. Let G_w denote the genus two Eisenstein series [12] of weight w, i.e., the genus two Siegel modular form of weight w defined by

$$G_w(\tau)=\sum\det(c\tau+d)^{-w} \qquad (w\geqq 4)$$

where the summation is extended over the bottom halves $(c\ \ d)$ of matrices M in $Sp(2,\mathbf{Z})$ which vary over distinct cosets of the subgroup of $Sp(2,\mathbf{Z})$ defined by $c=0$. Let g_w denote the classical (genus one) Eisenstein series of weight w, which is defined in precisely the same manner relative to the group $Sp(1,\mathbf{Z})$. Further, $\theta_{\mathfrak{m}}$ will denote the genus two theta-constant of characteristic $\mathfrak{m}$ [8], i.e., the holomorphic function defined in $\mathfrak{S}_2$ by

$$\theta_{\mathfrak{m}}(\tau)=\sum\exp\pi i\{{}^t(p+\tfrac{1}{2}m')\tau(p+\tfrac{1}{2}m')+{}^t(p+\tfrac{1}{2}m')m''\},$$

where the summation is taken over all integer column vectors p with two coordinates, and where $\mathfrak{m}=(m',m'')$ is a pair of integer column vectors m' and m'', each with two coordinates. We say that the characteristic $\mathfrak{m}$ is *even* if the scalar product ${}^tm'm''$ of m' and m'' is even. If $\mathfrak{m}$ is even, then $\theta_{\mathfrak{m}}$ does not vanish identically; and since $\theta_{\mathfrak{m}}$ is determined up to a sign by $\mathfrak{m}$ mod 2, we have ten non-zero theta-constants of genus two, whose product we denote by θ. The function θ^2 is modular form of weight 10 (cf. [9]).

The *diagonal plane* in $\mathfrak{S}_2$ is the image of the trivial imbedding ϕ_0 of

$(\mathfrak{S}_1)^2$ in $\mathfrak{S}_2$. Let (ϕ, Φ) be a fixed orthogonal modular imbedding for some real quadratic field K. If f is a holomorphic function in $\mathfrak{S}_2$, then by $f^{\#}$ we shall denote the restriction of f to the diagonal plane, i.e., $f^{\#} = f\phi_0$, and by f^* we shall denote the restriction of f to the image of ϕ, i.e., $f^* = f\phi$. Furthermore, if g is a holomorphic function in $(\mathfrak{S}_1)^2$, then by g^* we shall denote the restriction of g to the diagonal line, i.e., $g^*(\tau) = g(\tau, \tau)$. If f is a Siegel modular form (of genus two) of weight w, then f^* is a symmetric Hilbert modular form for K of weight w, and $f^{\#}$ is an elliptic modular form in each variable when the other variable is fixed (the additional condition required in the definition of modular forms of genus one is a consequence of the fact that the higher dimensional analogue of this condition is automatically satisfied by f, as is well known). Since the fixed modular imbedding is orthogonal, the elliptic modular forms f^{**} and $f^{\#*}$ are equal. Since $\theta^{\#} = 0$ (cf. [4]), we see that $\theta^{**} = 0$. Witt [13] has proved that $f^{\#}$ must be an element of $\mathcal{E}_w \otimes \mathcal{E}_w$ where $\mathcal{E}_w$ denotes the complex vector space of elliptic modular forms of weight w. In view of this and since $\dim \mathcal{E}_4 = \dim \mathcal{E}_6 = 1$, we see that $G_4^{**} = g_4^2$ and $G_6^{**} = g_6^2$. It is well known that g_4^2 and g_6^2 are algebraically independent over the field $\boldsymbol{C}$ of complex numbers. In the next section we shall prove that θ^* is not identically zero. As a result, the following theorem is obtained:

THEOREM 4.1. *If K is a real quadratic field whose discriminant is representable as the sum of two squares, then there are three algebraically independent symmetric Hilbert modular forms for K of weights four, six, and ten, namely G_4^*, G_6^*, and $(\theta^2)^*$.*

It is known that there is a symmetric Hilbert modular form h_2 of weight 2 such that $h_2^* = g_4$ (cf. [6], [10]). (Hence, h_2, G_6^*, and $(\theta^2)^*$ are algebraically independent.). Now the $*$ operator is a homomorphism of the graded ring $\mathfrak{H}^{(2)}$ of symmetric Hilbert modular forms of even weight to the graded ring $\mathcal{E}^{(4)}$ of elliptic modular forms whose weights are divisible by four. As Gundlach [3] has observed, $\mathcal{E}^{(4)}$ is the polynomial ring generated over $\boldsymbol{C}$ by g_4 and g_6^2. As a result, the $*$ homomorphism is onto $\mathcal{E}^{(4)}$, and one is led to investigate the kernel. In the case where this kernel is a principal ideal in $\mathfrak{H}^{(2)}$, one has the following lemma (cf. [3]):

LEMMA 4.2 (Gundlach). *Let h_2 and h_6 be symmetric Hilbert modular forms of weights two and six, respectively, such that $h_2^* = g_4$ and $h_6^* = g_6^2$. If h_0 is a symmetric Hilbert modular form of even weight which generates the ideal of all symmetric Hilbert modular forms h of even weight for which $h^* = 0$, then every symmetric Hilbert modular form of even weight can be written uniquely as a polynomial in h_2, h_6, and h_0.*

In the case of the real quadratic field of discriminant 5, Gundlach showed that $h_0 = (\theta^2)^*$ (defined without the use of a modular imbedding) satisfies the hypothesis of (4.2), and, consequently, $\mathfrak{A}^{(2)}$ is the polynomial ring generated over $\boldsymbol{C}$ by h_2, $G_6{}^*$, and $(\theta^2)^*$. We shall give a new proof of this in the next section.

5. The zeros of θ^*. Let (ϕ, Φ) be an orthogonal modular imbedding for the real quadratic field K of discriminant D. In this section we shall show that the zeros of θ^* lie on certain irreducible quadratic curves in $(\mathfrak{S}_1)^2$, and we shall study the equivalence of these curves under Hilbert modular mappings. Our basic tool is Igusa's result that every zero of θ in $\mathfrak{S}_2$ is equivalent to a point of the diagonal plane (see [4]). Thus, the zeros of θ lie on the fixed point varieties of matrices L in $Sp(2, \boldsymbol{Z})$ which are conjugate in $Sp(2, \boldsymbol{Z})$ to the matrix $H = \begin{pmatrix} h & 0 \\ 0 & h \end{pmatrix}$ where $h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Now Satake ([11], p. 488, Lemma 7) has shown that these matrices L are characterized as elements of $Sp(2, \boldsymbol{Z})$ by the conditions: (i) $L^2 = 1_4$, (ii) $L \neq 1_4$, and (iii) $L \equiv 1_4 \bmod 2$. The zeros of θ^* are precisely those points τ in $(\mathfrak{S}_1)^2$ for which $\phi(\tau)$ lies on the fixed point variety of a matrix L in $Sp(2, \boldsymbol{Z})$ which satisfies the conditions (i), (ii), and (iii). In order to formulate our results it is convenient to introduce the notion of *skew-Hermitian curve.*

We shall regard K as a subfield of the field of real numbers. A *skew-Hermitian matrix* over K is a 2-by-2 matrix A, with coefficients in K, which satisfies $A' = -{}^tA$, where A' is the conjugate matrix of A, obtained by conjugating each coefficient of A. With a skew-Hermitian matrix A we associate the quadratic curve in $\boldsymbol{C}^2$ defined by

$$ {}^t\begin{pmatrix} z_1 \\ 1 \end{pmatrix} A \begin{pmatrix} z_2 \\ 1 \end{pmatrix} = 0. $$

This curve meets $(\mathfrak{S}_1)^2$ if and only if the determinant of A is positive, and in this case the curve is uniquely determined by its intersection $C(A)$ with $(\mathfrak{S}_1)^2$. Thus, $C(A_1) = C(A_2)$ if and only if $A_2 = qA_1$ where q is a rational number. If μ is an element of $Sp(1, \mathfrak{O})$, then the inverse image $\mu^{-1}C(A)$ is equal to $C({}^t\mu A \mu')$. We shall say that A_1 and A_2 are *equivalent* if $A_2 = {}^t\mu A_1 \mu'$. Finally, let us observe that a skew-Hermitian matrix A with coefficients in $\mathfrak{O}$ can be written in the form $B - {}^tB'$, where B is a 2-by-2 matrix with coefficients in $\mathfrak{O}$, if and only if the elements of A on the principal diagonal are rational integral multiples of $D^{\frac{1}{2}}$. We see, therefore, that this latter property of A is preserved under equivalence.

PROPOSITION 5.1. *The zeros of* θ^* *lie on curves* $C(A)$ *defined by skew-Hermitian matrices* A, *with* $\det(A) > 0$, *which satisfy:*

(i) *There is a matrix* B *with coefficients in* $\mathfrak{O}$ *such that* $A = B - {}^tB'$.

(ii) *There is a rational integer* k *such that* $D - 4\det(A) = k^2$.

Proof. If a matrix L is conjugate to H in $Sp(2, \mathbf{Z})$, we shall show that the inverse image with respect to ϕ of the fixed point variety of L is such a curve $C(A)$. Now there is an orthogonal matrix

$$e = \begin{pmatrix} \cos\omega & \sin\omega \\ \sin\omega & -\cos\omega \end{pmatrix}$$

such that $N = \begin{pmatrix} e & 0 \\ 0 & e \end{pmatrix}$ defines the orthogonal modular imbedding (ϕ, Φ) as an alteration of (ϕ_0, Φ_0). Then $\phi(\tau)$ is a fixed point of L if and only if τ^* is a fixed point of the matrix $M = N^{-1}LN$. Suppose that

$$L = \begin{pmatrix} p & q \\ r & {}^tp \end{pmatrix} \text{ with } p = \begin{pmatrix} p_1 & p_2 \\ p_3 & -p_1 \end{pmatrix}\ q = \begin{pmatrix} 0 & q_2 \\ -q_2 & 0 \end{pmatrix}\ r = \begin{pmatrix} 0 & r_2 \\ -r_2 & 0 \end{pmatrix};$$

then the conditions on p_1, p_2, p_3, q_2, r_2 in order that L be conjugate to H are precisely

$$p_1^2 + p_2p_3 - q_2r_2 = 1; \qquad p_1 \equiv 1, p_2 \equiv p_3 \equiv q_2 \equiv r_2 \equiv 0 \bmod 2.$$

If $M = N^{-1}LN = \begin{pmatrix} a & b \\ c & {}^ta \end{pmatrix}$, and if we use notation for the coefficients of a, b, c which is similar to our notation for the coefficients of p, q, r, then we see that $b_2 = -q_2$, $c_2 = -r_2$, and

$$\begin{aligned} a_1 &= p_1(\cos^2\omega - \sin^2\omega) + (p_2 + p_3)\cos\omega\sin\omega, \\ a_2 &= 2p_1\cos\omega\sin\omega - p_2\cos^2\omega + p_3\sin^2\omega, \\ a_3 &= 2p_1\cos\omega\sin\omega + p_2\sin^2\omega - p_3\cos^2\omega. \end{aligned}$$

Since $\cos(2\omega) = uD^{-\frac{1}{2}}$ and $\sin(2\omega) = vD^{-\frac{1}{2}}$ (with u, v as in (3.4)), the matrix A defined by

$$A = \tfrac{1}{2}D^{\frac{1}{2}}\begin{pmatrix} c_2 & a_3 \\ -a_2 & -b_2 \end{pmatrix}$$

is a skew-Hermitian matrix which satisfies the condition (i). Furthermore, τ^* is a fixed point of M if and only if $\tau = (\tau_1, \tau_2)$ is a point of the curve $C(A)$. It is clear that $4 \cdot \det(A) = D(a_2a_3 - b_2c_2)$; and since $a_2a_3 - b_2c_2 = 1 - a_1^2$, we find that $4\det(A) = D - k^2$, where k is the rational integer $k = D^{\frac{1}{2}}a_1$.

The preceding proposition gives necessary but not sufficient conditions on a skew-Hermitian matrix A in order that it be obtained by the above procedure from an element L of $Sp(2,\boldsymbol{Z})$ which is conjugate to H. The determination of necessary and sufficient conditions does not seem to be useful although it can be carried out. However, it is of interest to see which integers k are obtained by the above procedure.

PROPOSITION 5.2. *For every integer k such that $k^2 < D$ and $k \equiv D$* mod 2, *there is a zero curve of θ^* defined by a skew-Hermitian matrix A which satisfies* $4\det(A) = D - k^2$.

Proof. We shall show that there is an odd integer p_1 and there are even integers p_2, p_3, q_2, r_2 such that $p_1{}^2 + p_2p_3 - q_2r_2 = 1$ and $k = up_1 + \frac{1}{2}v(p_2 + p_3)$. Once this is done, we let A be the skew-Hermitian matrix derived from the matrix L which is given by these numbers. Then $4\det(A) = D - k^2$, and the condition $k^2 < D$ implies that $\det(A)$ is positive. It is sufficient to find an odd integer p_1 and an even integer p_2 such that $k = up_1 + \frac{1}{2}vp_2$. Using the fact that D is the discriminant of a real quadratic field, we see that the integers u and $\frac{1}{2}v$ are coprime. Since $k \equiv D \equiv u \bmod 2$, p_1 and p_2 may be so chosen.

COROLLARY 5.3. *If every zero of θ^* is congruent* $\operatorname{mod} Sp(1, \mathfrak{O})$ *to a diagonal point of $(\mathfrak{S}_1)^2$, then the discriminant D is* 5.

Proof. To be completely explicit we should say that by *diagonal point* of $(\mathfrak{S}_1)^2$ we mean a point $\tau = (\tau_1, \tau_2)$ for which $\tau_1 = \tau_2$. The diagonal points are precisely the locus $C(j)$ given by the skew-Hermitian matrix $j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Let k be an integer which satisfies the hypothesis of (5.2), and choose a skew-Hermitian matrix A such that $4\det(A) = D - k^2$ and $C(A)$ is a zero curve of θ^*. Then every point of $C(A)$ is equivalent to a diagonal point. Since $C(A)$ is uncountably infinite while $Sp(1, \mathfrak{O})$ is countably infinite, we can find an element μ in $Sp(1, \mathfrak{O})$ such that $C(A) = \mu^{-1}C(j)$, and from this it follows that $A = q^t\mu j \mu'$. Since $\det(A) = q^2$ is a rational integer, q is also a rational integer, and $4\det(A) = D - k^2$ is a perfect square. Thus, for every integer k with $k^2 < D$ and $k \equiv D \bmod 2$, the integer $D - k^2$ must be a perfect square. The only positive integers D which satisfy this condition are 1, 4, and 5, and, of these, only 5 is a discriminant. Next we shall prove some lemmas which are helpful in determining the number of inequivalent zero curves of θ^* for small values of D.

LEMMA 5.4. *If the ideal class number of K is one, then a skew-*

Hermitian matrix A with coefficients in $\mathfrak{O}$ is equivalent to one of the form $\begin{pmatrix} 0 & * \\ * & * \end{pmatrix}$ *if (and only if)* $\det(A)$ *is the norm of some element of K.*

Proof. One can always find an element of $Sp(1,\mathfrak{O})$ whose first column is a preassigned pair of coprime integers. Thus it suffices to find such a column ζ for which ${}^t\zeta A\zeta' = 0$. Since $\mathfrak{O}$ is a principal ideal domain, we see that this is equivalent to the existence of an element λ of K for which (λ, λ') lies on the quadratic curve defined by A. If $A = \begin{pmatrix} \rho & \sigma \\ -\sigma' & \eta \end{pmatrix}$, then this latter condition can be put in the form $Nm(\rho\lambda - \sigma') = \det(A)$ whenever $\rho \neq 0$. Clearly, this equation has a solution λ in K if and only if $\det(A)$ is a norm.

LEMMA 5.5. *If the ideal class number of K is one, then the norm of the fundamental unit of K is -1.*

Proof. Since D is the sum of two squares, there is an ideal $\mathfrak{A}$ such that the narrow ideal class of $\mathfrak{A}^2$ is the narrow ideal class of the different of K. Since $\mathfrak{A}$ is principal, the narrow ideal class of $\mathfrak{A}^2$ is also equal to the narrow ideal class of $\mathfrak{O}$. This can happen if and only if the norm of the fundamental unit is -1.

LEMMA 5.6. *If the ideal class number of K is one, then a skew-Hermitian matrix $A = B - {}^tB'$, where the coefficients of B lie in $\mathfrak{O}$, is equivalent to $\pm j$, where $j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, if (and only if)* $\det(A) = 1$.

Proof. From (5.4) we may assume that $A = \begin{pmatrix} 0 & \sigma \\ -\sigma' & \eta \end{pmatrix}$, where σ is a unit in $\mathfrak{O}$ of norm 1. Since $A = B - {}^tB'$, we have $\eta = \nu - \nu'$ for some element ν of $\mathfrak{O}$. Let $\lambda = \nu'\sigma^{-1}$, and $\mu = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$. Then ${}^t\mu A\mu' = \begin{pmatrix} 0 & \sigma \\ -\sigma' & 0 \end{pmatrix}$, and we shall denote this new matrix by A. Since σ has norm 1, and the fundamental unit of K has norm -1, there is a unit σ_1 in $\mathfrak{O}$ such that $\sigma = \pm\sigma_1{}^2$. If μ now stands for

$$\begin{pmatrix} \sigma_1{}^{-1} & 0 \\ 0 & \sigma_1 \end{pmatrix},$$

then ${}^t\mu A\mu' = \pm j$.

LEMMA 5.7. *Suppose that the ideal class number of K is one and that A is a skew-Hermitian matrix which satisfies the conditions* (i) *and* (ii) *of* (5.1). *If* $\det(A)$ *is a rational prime, then* $\det(A)$ *is the norm of some (prime) integer of K; and if ρ denotes a given (prime) integer of K whose norm is* $\det(A)$, *then A is equivalent to* $\pm\begin{pmatrix} 0 & \rho \\ -\rho' & 0 \end{pmatrix}$.

Proof. Since the norm of the fundamental unit is -1, for the first assertion it suffices to observe that $-\det(A)$ is the norm of the integer $\frac{1}{2}(k+D^{\frac{1}{2}})$. By (5.4) we may assume that $A = \begin{pmatrix} 0 & \sigma \\ -\sigma' & \eta \end{pmatrix}$. We may write, as in (5.6), $\eta = \nu - \nu'$. Since the norm of σ is a prime, every element of $\mathfrak{O}$ is congruent $\operatorname{mod} \sigma$ to a rational integer. Therefore, we may assume that $\nu = \lambda\sigma$ for some element λ of $\mathfrak{O}$. Then $A = {}^t\mu A_0 \mu'$, where $A_0 = \begin{pmatrix} 0 & \sigma \\ -\sigma' & 0 \end{pmatrix}$ and $\mu = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$. Either σ or σ' differs by a unit from ρ. We may assume that σ differs by a unit from ρ after replacing A_0, if necessary, by ${}^t j A_0 j$. If $\sigma = \rho\epsilon$, then ϵ is a unit of norm 1, and we have $\epsilon = \pm \epsilon_1{}^2$ for some unit ϵ_1. Taking μ as in (5.6), we have the desired result.

If we let K denote the real quadratic field of discriminant 5, then from (5.1) we see that $\det(A) = 1$ for every zero curve of θ^*. From (5.6) we see that every zero of θ^* is equivalent to a diagonal point. This is the crucial point in the proof of the structure theorem of Gundlach for the graded ring of symmetric Hilbert modular forms of even weight. For by (4.2) it suffices to show that if h is a symmetric Hilbert modular form of even weight for which $h^* = 0$, then h is divisible by $(\theta)^2$. From the "Fourier expansion" of θ (see [9]), it is plain that θ vanishes to the first order on the diagonal plane in $\mathfrak{S}_2$. It follows easily that θ^* vanishes to the first order along the diagonal line in $(\mathfrak{S}_1)^2$. Thus, if h is given as above, then θ^* divides h, say, $h = \theta^* h_1$. Since θ^* is "skew-symmetric," one sees that h_1 is skew-symmetric, and, therefore, $h_1{}^* = 0$. From this it is clear that θ^* divides h_1, which is the desired result.

6. The quadratic field of discriminant 8. The first few discriminants for which modular imbeddings exist are 5, 8, 13, and 17. We know how to obtain the structure of the graded ring of symmetric Hilbert modular forms of even weight when the ideal of such modular forms which vanish on the diagonal line in $(\mathfrak{S}_1)^2$ is a principal ideal, and we know that this is indeed the case when the discriminant $D = 5$. We have seen that $(\theta^*)^2$ generates the ideal in that case, and we know that in every other case $(\theta^*)^2$ has additional zeros so that it seems to be a rather unlikely candidate for generator. On the other hand, if the said ideal is principal (which need not be the case), then the generator must be a factor of $(\theta^*)^2$, and in this way one is led to look for such factors. The most obvious way that one could try to do this is by forming partial products of the $\theta_m{}^*$. This turns out to be of limited interest because it breaks down completely in the case $D = 13$ due to the fact that the image under Φ of $Sp(1,\mathfrak{O})$ operates transitively on the even characteristics (see [8]). In this section we shall employ the "partial product" idea successfully in the case $D=8$.

At first, we shall let K denote the real quadratic field of arbitrary *even* discriminant D. Let (ϕ,Φ) be an orthogonal modular imbedding for K which corresponds in the sense of (3.4) to a pair of integers (u,v). We have already observed, in the proof of (5.2) that the integers u and $\frac{1}{2}v$ are coprime. Using the equation (3.7) for the image of ϕ, we see that ϕ maps the ring $\mathfrak{O}$ into the set of all 2-by-2 integer matrices $n=\begin{pmatrix} n_1 & n_{12} \\ n_{21} & n_2 \end{pmatrix}$ which satisfy the conditions $n_1 \equiv n_2$ and $n_{12} \equiv n_{21} \bmod 2$. These conditions are the same as ${}^tjnj \equiv n \bmod 2$, where $j=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Since j is orthogonal, this set forms a ring $R(j)$ which has the additional property that if n is in $R(j)$ then ${}^tn$ is also in $R(j)$. It is obvious that the image of Φ is the set of $M=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in $Sp(2,\mathbf{Z})$ such that a, b, c, d are in the image of ϕ, and, therefore, the image of Φ is contained in the set of M for which a, b, c, d are elements of the ring $R(j)$. In other words, the matrix M satisfies the condition ${}^tJMJ \equiv M \bmod 2$, where $J=\begin{pmatrix} j & 0 \\ 0 & j \end{pmatrix}$. This set of M forms a subgroup $Sp(2,\mathbf{Z},j)$ of $Sp(2,\mathbf{Z})$ which contains the "principal congruence subgroup of level two," i.e., the subgroup of $Sp(2,\mathbf{Z})$ defined by $M \equiv 1_4 \bmod 2$. Since the principal congruence subgroup has finite index in $Sp(2,\mathbf{Z})$, it follows that $Sp(2,\mathbf{Z},j)$ has finite index in $Sp(2,\mathbf{Z})$. Since the image of $Sp(1,\mathfrak{O})$ is contained in $Sp(2,\mathbf{Z},j)$, if f is a holomorphic function in $\mathfrak{S}_2$ which satisfies the functional equation for a modular form of weight w with respect to the operations in $\mathfrak{S}_2$ of the elements of $Sp(2,\mathbf{Z},j)$, then $f^* = f\phi$ is a symmetric Hilbert modular form of weight w.

An integer column vector with two components will be called *mod 2 diagonal* if its two components are congruent mod 2. A column vector m is mod 2 diagonal if and only if $jm \equiv m \bmod 2$. Thus, if m is a mod 2 diagonal column vector and if n is an element of the 2-by-2 matrix ring $R(j)$, then nm is a mod 2 diagonal column vector. A characteristic $\mathfrak{m}=(m',m'')$ will be called *mod 2 diagonal* if the column vectors m' and m'' are mod 2 diagonal. The characteristic $\mathfrak{m}$ is, then, mod 2 diagonal if and only if it is fixed by the operation of J on the set of characteristics mod 2 (see [8]). It follows that the set of mod 2 diagonal characteristics is stable under the operations of the group $Sp(2,\mathbf{Z},j)$. From the "transformation law of theta-constants" in [8] it is obvious that θ_2, the product of the four theta-constants

$\theta_{\mathfrak{m}}$ for which $\mathfrak{m}$ is mod 2 diagonal, possesses a functional equation with respect to the mappings in $\mathfrak{S}_2$ of the group $Sp(2,\mathbf{Z},j)$. Without much difficulty one can prove, using the transformation law, that

$$\theta_2(M\cdot\tau)^2=\det(c\tau+d)^4\theta_2(\tau)^2 \qquad (M \text{ in } Sp(2,\mathbf{Z},j))$$

and that $\theta_2(J\cdot\tau)=-\theta_2(\tau)$. As a result, $(\theta_2{}^*)^2$ is a symmetric Hilbert modular form of weight 4, and $\theta_2{}^*$ changes its sign under the permutation of the coordinates in $(\mathfrak{S}_1)^2$. In ths case $D=8$ we shall show that $(\theta_2{}^*)^2$ generates the ideal of all symmetric Hilbert modular forms of even weight which vanish on the diagonal line.

We now restrict to the case $D=8$. Let $C(A)$ be an arbitrary zero curve of θ^*. According to (5.1) $\det(A)$ is either 1 or 2, and according to (5.2) there are zero curves of θ^* corresponding to both of these values for $\det(A)$. From (5.6), if $\det(A)$ is 1, then $C(A)$ is equivalent to the diagonal line $C(j)$. Let $\epsilon=1+2^{\frac{1}{2}}$, the fundamental unit of K. Then $\epsilon 2^{\frac{1}{2}}$ is a prime element in $\mathfrak{O}$ of norm 2, and by (5.7) every $C(A)$ for which $\det(A)=2$ is equivalent to the line

$$\epsilon\tau_1+\epsilon'\tau_2=0. \tag{6.1}$$

Let us take the orthogonal modular imbedding corresponding, in the sense of (3.4), to $(u,v)=(2,2)$. In the notation of Section 5, consider the matrix L whose components p, q, r are given by $p=\begin{pmatrix}1 & 0\\ -2 & -1\end{pmatrix}$, $q=r=0$. The inverse image with respect to ϕ of the fixed point variety of L is the line (6.1). It is not difficult to see that $\theta_3=\theta/\theta_2$ changes its sign under the mappings in $\mathfrak{S}_2$ defined by L. This means that θ_3 vanishes on the fixed point variety of L, and, therefore, $\theta_3{}^*$ equals zero on the line (6.1). It is a consequence of Igusa's result on the zeros of θ, which we mentioned in the beginning of Section 5, that at most one theta-constant (of even characteristic) can vanish at a given point of $\mathfrak{S}_2$. This means that $\theta_2{}^*$ does not vanish on the line (6.1) nor on any curve which is equivalent to it. Hence, *every zero of* $\theta_2{}^*$ *is equivalent to a point of the diagonal line.* Copying the argument at the end of Section 5, we easily show that the ideal of symmetric Hilbert modular forms of even weight which vanish on the diagonal line is the principal ideal generated by $(\theta_2{}^*)^2$. Applying (4.2) we obtain:

THEOREM 6.2. *Every symmetric Hilbert modular form of even weight for the real quadratic field of discriminant* 8 *can be written uniquely as a polynomial in* h_2, $G_6{}^*$, *and* $(\theta_2{}^*)^2$.

Added in proof. We find that we have not given an adequate explanation of the statement in Section 2 that the condition (iii) of Definition 2.1 can be replaced by the condition (iv). It is not obvious that (iv) is a consequence of (iii). Let us denote by (v) the condition of Proposition 2.2. It is obvious that (v) implies (iv) and that (iv) implies (iii). We shall now show that (iii) implies (v). The main point is that there is a positive integer w such that the Siegel modular forms of weight w have no common zero. This is a consequence of either (1) the compactification theory of Satake and Baily or (2) the theory of theta-constants [8]. As a result, we see that

$$R(\Phi\mu, \phi\tau)^w = R(\mu^*, \tau^*)^w,$$

and, therefore, the function h, which appears in the proof of Proposition 2.2, has the property that h^w is a modular invariant. Thus, h^w is constant, and so h is constant. One now proves that (v) is satisfied by using the argument already given for Proposition 2.2.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] O. Blumenthal, "Uber Modulfunktionen von mehreren Veränderlichen," *Mathematische Annalen*, vol. 56 (1903), pp. 509-548, and vol. 58 (1904), pp. 497-527.

[2] L. E. Dickson, *History of the Theory of Numbers*, vol. III, Washington, 1923.

[3] K.-B. Gundlach, "Die Bestimmung der Funktionen zur Hilbertschen Modulgruppe des Zahlkörpers $Q(\sqrt{5})$," *Mathematische Annalen*, vol. 152 (1963), pp. 226-256.

[4] W. F. Hammond, "On the graded ring of Siegel modular forms of genus two," *American Journal of Mathematics*, vol. 87 (1965), pp. 502-506.

[5] E. Hecke, *Vorlesungen uber die Theorie der algebraischen Zahlen*, Leipzig, 1923.

[6] ———, "Analytische Funktionen und algebraische Zahlen II," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 3 (1924), pp. 213-236.

[7] J. Igusa, "On Siegel modular forms of genus two," *American Journal of Mathematics*, vol. 84 (1962), pp. 175-200.

[8] ———, "On the graded ring of theta-constants," *American Journal of Mathematics*, vol. 86 (1964), pp. 219-246.

[9] ———, "On Siegel modular forms of genus two (II)," *American Journal of Mathematics*, vol. 86 (1964), pp. 392-412.

[10] H. D. Kloosterman, "Theorie der Eisensteinschen Reihen von mehreren Veränderlichen," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 6 (1928), pp. 163-188.

SIMULTANEOUS APPROXIMATION TO A BASIS OF A REAL NUMBERFIELD.

By Wolfgang M. Schmidt.

---

1. Introduction. Let $\psi(q) > 0$ be a decreasing function of positive integers q, such that

$$\Psi(h) = \sum_{q=1}^{h} \psi(q)^n$$

tends to infinity. Write $N(h; \alpha_1, \cdots, \alpha_n)$ for the number of $(n+1)$-tuples of integers $q; p_1, \cdots, p_n$ with $1 \leqq q \leqq h$ and

$$0 < \alpha_i q + p_i \leqq \psi(q) \qquad (1 \leqq i \leqq n).$$

It is known [1, 4] that $N(h; \alpha_1, \cdots, \alpha_n)$ is for almost all $(\alpha_1, \cdots, \alpha_n)$ in the sense of Lebesgue asymptotically equal to $\Psi(h)$ for large h. Hence one would expect this to be true of many given n-tuples, say of $(\pi^1, \cdots, \pi^n)$. One would be more hesitant to expect this of $(\beta_1, \cdots, \beta_n)$ if $1, \beta_1, \cdots, \beta_n$ is a basis of a real numberfield of degree $n+1$, because of the exceptional status of such n-tuples in diophantine approximations. In fact, $N(h; \beta_1, \cdots, \beta_n) = 0$ for all h if $\psi(q) = cq^{-1/n}$ and $0 < c < c_0(\beta_1, \cdots, \beta_n)$. But the asymptotic formula does hold under a further mild assumption on $\psi(q)$:

Theorem 1. *Let $\psi(q)$, $\Psi(h)$ be as before, and assume that*

$$\omega(q) = q\psi(q)^n$$

tends to infinity. Let $1, \beta_1, \cdots, \beta_n$ be a basis of a real algebraic numberfield. Then

$$N(h; \beta_1, \cdots, \beta_n) = \Psi(h) + O\Big(\sum_{q=1}^{h} \psi(q)^n \omega(2q)^{-1/(n+1)}\Big)$$
$$= \Psi(h)(1 + o(1)).$$

There is a dual theorem. Let $\phi(q) > 0$ be a decreasing function of positive integers q, such that

$$\chi(q) = q^n \phi(q)$$

tends to infinity. Put

Received May 17, 1965.

$$\Phi(h) = \sum \phi(\max(q_1, \cdots, q_n))$$

where the summation is over all integral n-tuples $q_1, \cdots, q_n$ with $1 \leqq q_i \leqq h$ $(i = 1, \cdots, n)$. $\Phi(h)$ tends to infinity and is asymptotically equal to

$$n \sum_{q=1}^{h} q^{n-1} \phi(q).$$

Put $M(h; \alpha_1, \cdots, \alpha_n)$ for the number of integral $(n+1)$-tuples $q_1, \cdots, q_n; p$ with $1 \leqq q_i \leqq h$ $(i = 1, \cdots, n)$ and

$$0 < \alpha_1 q_1 + \cdots + \alpha_n q_n + p \leqq \phi(\max(q_1, \cdots, q_n)).$$

THEOREM 2. *Let $\phi(q)$, $\Phi(h)$ be as above, and let $1, \beta_1, \cdots, \beta_n$ be a basis of a real numberfield. Then*

$$M(h; \beta_1, \cdots, \beta_n) = \Phi(h) + O\Big(\sum_{q=1}^{h} q^{n-1} \phi(q) \chi(2q)^{-1/(n+1)}\Big)$$
$$= \Phi(h)(1 + o(1)).$$

We shall prove Theorem 1 only; Theorem 2 can be proved similarly. The idea of our proof is to use the units of numberfields to construct automorphisms of certain point-lattices.

I am indebted to S. Lang who sent me his manuscript [3] before publication. He was the first to prove a result of the type of our theorems. He showed for quadratic irrationals β and under further restrictions on $\psi(q)$ that $N(h; \beta)$ is asymptotically equal to $c(\beta)\Psi(h)$, where $c(\beta)$ was a constant depending on β. He took the inequality $|\beta q + p| \leqq \psi(q)$ rather than our $0 < \beta q + p \leqq \psi(q)$, and hence our results show $c(\beta) = 2$. Actually, Lang obtained $c(\beta) = c_2(\beta) c_L(\beta)$ where c_2, c_L were certain well defined functions of β, and a very short computation gives $c(\beta) = 2$. His related result [2] probably cannot easily be generalized to simultaneous approximations.[1]

2. Remarks on number fields. Write $\mathbf{Z}$ for the ring of rational integers, $\boldsymbol{Q}$ for the field of rationals. Let F be a real number field of degree

$$m = n + 1. \tag{2.1}$$

Let $\tau^{(0)} =$ identity, $\tau^{(1)}, \cdots, \tau^{(n)}$ be the isomorphisms of F into the field of complex numbers. Write $F^{(i)}=\tau^{(i)}F$, and for $\xi\in F$ put $\xi^{(i)}=\tau^{(i)}\xi$. There will be integers r, s with $r+2s=m$, such that r of the fields $F^{(i)}$ are real, and $2s$ of them complex. We may assume $F^{(0)},\cdots,F^{(r-1)}$ to be real, and $\xi^{(r+s+j)}$ to be the complex conjugate of $\xi^{(r+j)}$ for every $\xi\in F$ $(0\leqq j\leqq s-1)$. We have $m\geqq 2$,$r\geqq 1$, hence

---

[1] *Note added to proofs.* S. Lang in a paper "asymptotic diophantine approximation" (*Proceedings of the National Academy of Sciences*, vol. 55 (1966), pp. 31-34) gave a short proof of a result which includes our case $n = 1$, $\omega(q)$ increasing. He admits not only quadratic irrationals but general classes of badly approximable numbers.

$$t=r+s\geqq 2.$$

Put

$$\xi^{*(i)}=\begin{cases}\xi^{(i)} \text{ if } 0\leqq i\leqq r-1,\\ \operatorname{Re}(\xi^{(i)}) \text{ if } r\leqq i\leqq r+s-1,\\ \operatorname{Im}(\xi^{(i)}) \text{ if } r+s\leqq i\leqq n.\end{cases}$$

Here Re, Im denote the real and imaginary parts.

By a *module in* F we mean a $\boldsymbol{Z}$-module R consisting of elements $q_0\gamma_0+\cdots+q_n\gamma_n$ with $q_i\in\boldsymbol{Z}$ $(0\leqq i\leqq n)$ and $\gamma_0,\cdots,\gamma_n$ a fixed basis of F over $\boldsymbol{Q}$. With R we associate the point-lattice $\Lambda(R)$ consisting of all points

$$(\xi^{*(0)},\cdots,\xi^{*(n)}) \qquad (\xi\in R)$$

in Euclidean space E_m. Write $\Delta(R)$ for the determinant of $\Lambda(R)$. $\Delta(R)$ is the absolute value of the determinant of the matrix $(\gamma^*{}_j{}^{(i)})$ $(0\leqq i,j\leqq n)$. If R' is a submodule of R, $\Delta(R')/\Delta(R)$ is an integer. Given an integer $t>0$ and a module R there are only finitely many submodules R' having $\Delta(R')=t\Delta(R)$. If $\theta\in F$ is not zero, let θR be the module of all elements $\theta\xi$ with $\xi\in R$. $\Delta(\theta R)=|\theta^{(0)}\cdots\theta^{(n)}|\Delta(R)$.

Let $c_0,\cdots,c_n$ be reals having $c_i\neq 0$ $(0\leqq i\leqq r-1)$, $c_i^2+c_{i+s}^2\neq 0$ $(r\leqq i\leqq r+s-1)$. Write $(c_0,\cdots,c_n)\Lambda$ for the lattice of points $(\mu_0,\cdots,\mu_n)$ where

$$\mu_i=\begin{cases}c_i\lambda_i \text{ if } 0\leqq i\leqq r-1\\ c_i\lambda_i-c_{i+s}\lambda_{i+s} \text{ if } r\leqq i\leqq r+s-1\\ c_i\lambda_{i-s}+c_{i-s}\lambda_i \text{ if } r+s\leqq i\leqq n\end{cases}$$

and where $(\lambda_0,\cdots,\lambda_n)\in\Lambda$. In particular, for $\theta\neq 0$ in F put

$$\theta\Lambda=(\theta^{*(0)},\cdots,\theta^{*(n)})\Lambda.$$

Then

$$\Lambda(\theta R)=\theta\Lambda(R) \tag{2.2}$$

Let R_0 be the ring of algebraic integers in F and η a unit of R_0. Let R be an arbitrary module in F. There is a $d\in\boldsymbol{Z}$ with $dR\subset R_0$. $\Delta(dR)$

$= t\Delta(R_0)$ with $t \in \boldsymbol{Z}$. For every $k \in \boldsymbol{Z}$, $\eta^k dR \subset R_0$ and $\Delta(\eta^k dR) = \Delta(dR) = t\Delta(R_0)$. Thus there are integers $k_1 \neq k_2$ having $\eta^{k_1} dR = \eta^{k_2} dR$, and there is an integer $k_0 \neq 0$ with $\eta^{k_0} dR = dR$, $\eta^{k_0} R = R$. Putting $h = 2k_0$ we see: *There is a* $\theta = \eta^h$ *with* $h \neq 0$, $h \in \boldsymbol{Z}$ *such that* $\theta^{(i)} > 0$ $(0 \leqq i \leqq r-1)$ *and* $\theta R = R$.

Now in R there are $t-1$ units $\eta_1, \cdots, \eta_{t-1}$ such that the determinant $|\log|\eta_j^{(i)}||$ $(1 \leqq i, j \leqq t-1)$ is not zero. To each η_j we construct a corresponding θ_j. Using (2.2) we obtain

LEMMA 1. *Let R be a module in F, $\Lambda = \Lambda(R)$. There are elements $\theta_1, \cdots, \theta_{t-1}$ in F with $\theta_j^{(i)} > 0$ $(0 \leqq i \leqq r-1, 1 \leqq j \leqq t-1)$, with the determinant*

$$|\log|\theta_j^{(i)}|| \qquad (1 \leqq i, j \leqq t-1) \tag{2.3}$$

not zero and

$$\theta_i \Lambda = \Lambda \qquad (1 \leqq i \leqq t-1). \tag{2.4}$$

LEMMA 2. *Let F have basis $\beta_0, \beta_1, \cdots, \beta_n$ over $\boldsymbol{Q}$. There is a basis $\alpha_0, \alpha_1, \cdots, \alpha_n$ of F over $\boldsymbol{Q}$ having*

$$\alpha_0^{(i)}\beta_0 + \cdots + \alpha_n^{(i)}\beta_n = 0 \qquad (1 \leqq i \leqq n). \tag{2.5}$$

The determinant

$$\delta = |\alpha_j^{(i)}| \qquad (1 \leqq i, j \leqq n) \tag{2.6}$$

is not zero.

Proof. The matrix B with row-vectors $(\beta_0^{(i)}, \cdots, \beta_n^{(i)})$ $(0 \leqq i \leqq n)$ has rank m. The matrix B' obtained by deleting the first row of B has rank n. Put d_k $(k = 0, 1, \cdots, n)$ for $(-1)^k$ times the determinant of the $n \times n$-matrix obtained from B' by removing the k-th column. For some l, $d_l \neq 0$. Put $\alpha_0 = d_0/d_l, \cdots, \alpha_n = d_n/d_l$. Let N be the field of all rational expressions of elements of the fields $F^{(0)}, \cdots, F^{(n)}$. N is the least normal extension over $\boldsymbol{Q}$ of each field $F^{(i)}$ $(i = l, \cdots, n)$. An automorphism of N over $F^{(0)}$ will permute the rows of B' and hence will leave $\alpha_0, \cdots \alpha_n$ fixed. Thus $\alpha_0, \cdots, \alpha_n$ are in $F^{(0)} = F$. One has

$$\alpha_0\beta_0^{(i)} + \cdots + \alpha_n\beta_n^{(i)} = 0 \qquad (1 \leqq i \leqq n). \tag{2.7}$$

$\tau^{(j)}$ is an isomorphism of $F = F^{(0)}$ into $F^{(j)}$. It can be extended to an automorphism $\tau^{*(j)}$ of N. $\tau^{*(j)}$ will give a permutation of the fields $F^{(0)}, \cdots, F^{(n)}$, and since $F^{(0)}$ goes into $F^{(j)}$, the fields $F^{(1)}, \cdots, F^{(n)}$ will in some order go into $F^{(0)},\cdots,F^{(j-1)}, F^{(j+1)},\cdots,F^{(n)}$. Applying $\tau^{*(j)}$ to (2.7) we obtain

$$\alpha_0^{(j)}\beta_0^{(i)}+\cdots+\alpha_n^{(j)}\beta_n^{(i)}=0 \qquad (i\neq j). \tag{2.8}$$

(2.8) is true for any $i\neq j$ with $0\leqq i,j\leqq n$. In particular (2.5) holds.

Since B has rank m, since $\alpha_0,\cdots,\alpha_n$ are not all zero and because of (2.7), $\alpha_0\beta_0^{(0)}+\cdots+\alpha_n\beta_0^{(0)}\neq 0$. Thus

$$\alpha_0^{(i)}\beta_0^{(i)}+\cdots+\alpha_n^{(i)}\beta_n^{(i)}\neq 0 \qquad (0\leqq i\leqq n). \tag{2.9}$$

Let A be the matrix with row-vectors $(\alpha_j^{(0)},\cdots,\alpha_j^{(n)})$ $(0\leqq j\leqq \mathrm{n})$. By (2.8) the product BA is a diagonal matrix, and by (2.9) the elements in the diagonal are not zero. Thus BA and hence A is not singular, and $\alpha_0,\cdots,\alpha_n$ form a basis of F.

By an argument used before one sees that $\delta^2\in F^{(0)}=F$. Put $\delta^2=\mu$. Since $\mu^{(0)},\cdots,\mu^{(n)}$ are the squares of the determinants of certain $n\times n$-submatrices of the nonsingular matrix A, not all of them can be zero. Since they are conjugates, none of them is zero. In particular $\delta\neq 0$.

3. The number of points of $\mathbf{\Lambda}(R)$ in certain regions.

Let

$$L_i=L_i(x_1,\cdots,x_n)=\sum_{k=1}^{n} c_{ik}x_k \qquad (1\leqq i\leqq n)$$

be n real linear forms in n variables. Assume the absolute value a of the determinant $|c_{ik}|$ $(1\leqq i,k\leqq n)$ to be positive.

Let $\rho(x)>0$ be defined for $x\geqq 0$, decreasing and such that $\sigma(x)=x\rho^n(x)$ tends to infinity when x does.

LEMMA 3. *Let F be a real numberfield of degree $m=n+1$, R a module in F, $\Lambda=\Lambda(R)$. Let $L_1,\cdots,L_n$; ρ be as described above, and let $b>1$. Write $N(h)$ for the number of latticepoints $(\lambda_0,\cdots,\lambda_n)$ of Λ satisfying*

$$0<\lambda_0\leqq h,$$
$$0<\lambda_0(L_i(\lambda_1,\cdots,\lambda_n))^n\leqq\sigma(\lambda_0) \qquad (1\leqq i\leqq n).$$

Then

$$N(h)=\Delta(R)^{-1}a^{-1}\int_1^h x^{-1}\sigma(x)\,dx+0\Bigl(\int_1^h x^{-1}\sigma(x)\sigma(bx)^{-1/m}dx\Bigr).$$

Proof. $C_1, C_2,\cdots$ will be positive constants depending only on $L_1,\cdots,L_n, R,\rho,b$ and the choice of $\theta_1,\cdots,\theta_{t-1}$ in Lemma 1. Put $f=\sqrt{b}$. Write

$$M(k) = N(k) - N(k/f).$$

Suppose we can prove that for $k \geqq k_0 > f^2$

$$(3.1) \qquad | M(k) - \Delta(R)^{-1}a^{-1} \int_{k/f}^{k} x^{-1}\sigma(x)\,dx \,| \leqq C_1 \int_{k/f^2}^{k/f} x^{-1}\sigma(x)\sigma(bx)^{-1/m}dx.$$

Put l for the integral part of ${}_f\log(h/k_0)$, where ${}_f\log$ means the logarithm to the base f. We obtain for $h \geqq k_0 f$

$$\begin{aligned} | N(h) - \Delta(R)^{-1}a^{-1} \int_1^h x^{-1}\sigma(x)\,dx \,| & \\ \leqq \sum_{w=0}^{l-1} | M(hf^{-w}) - \Delta(R)^{-1}a^{-1} \int_{hf^{-w-1}}^{hf^{-w}} x^{-1}\sigma(x)\,dx \,| & \\ + N(k_0 f) + \Delta(R)^{-1}a^{-1} \int_1^{k_0 f} x^{-1}\sigma(x)\,dx & \\ \leqq C_1 \int_1^h x^{-1}\sigma(x)\sigma(bx)^{-1/m}dx + 0(1). & \end{aligned}$$

Hence it will suffice to prove (3.1).

Define $\mu_1, \cdots, \mu_{t-1}$ by

$$\sigma(k)^{1/nm}k^{-1/n} = | \theta_1^{(i)} |^{\mu_1} \cdots | \theta_{t-1}^{(i)} |^{\mu_{t-1}} \qquad (1 \leqq i \leqq t-1).$$

This is possible since the determinant (2.3) is not zero. Let $m_1, \cdots, m_{t-1}$ be integers with $| m_i - \mu_i | \leqq \frac{1}{2}$, and put $\kappa = \theta_1^{m_1} \cdots \theta_{t-1}^{m_{t-1}}$. There is a $C_2 > 0$ such that

$$C_2^{-1}\sigma(k)^{1/nm}k^{-1/n} \leqq | \kappa^{(i)} | \leqq C_2 \sigma(k)^{1/nm}k^{-1/n} \qquad (1 \leqq i \leqq n),$$

whence

$$C_2^{-n}\sigma(k)^{-1/m}k \leqq \kappa^{(0)} \leqq C_2^{n}\sigma(k)^{-1/m}k.$$

Define $c_0, \cdots, c_n$ by

$$\kappa^{*(i)} = \sigma(k)^{1/nm}k^{-1/n}c_i \qquad (1 \leqq i \leqq n),$$

$$\kappa^{*(0)} = \sigma(k)^{-1/m}kc_0.$$

One has $| c_i | \leqq C_2^n$ $(0 \leqq i \leqq n)$ and

$$(3.2) \qquad c_0 \cdots c_{r-1}(c_r^2 + c_{r+s}^2) \cdots (c_{r+s-1}^2 + c_n^2) = 1.$$

Let Λ' be the lattice

$$\Lambda' = (c_0, \cdots, c_n)\Lambda,$$

where $\Lambda = \Lambda(R)$. Then $\det(\Lambda') = \det(\Lambda)$ by (3.2).

$M(k)$ is the number of latticepoints $(\lambda_0, \cdots, \lambda_n)$ of Λ satisfying

$$k/f < \lambda_0 \leqq k, \tag{3.3}$$

$$0 < \lambda_0 (L_i(\lambda_1, \cdots, \lambda_n))^n \leqq \sigma(\lambda_0) \qquad (1 \leqq i \leqq n). \tag{3.4}$$

Now $\Lambda = \kappa\Lambda$, and therefore $M(k)$ is also the number of latticepoints $(\lambda_0, \cdots, \lambda_n)$ of $\kappa\Lambda$ satisfying (3.3) and (3.4). But

$$(\kappa^{*(0)}, \cdots, \kappa^{*(n)}) = (\sigma(k)^{-1/m} k c_0, \sigma(k)^{1/nm} k^{-1/n} c_1, \cdots, \sigma(k)^{1/nm} k^{-1/n} c_n),$$

hence $\kappa\Lambda$ consists of points

$$(\sigma(k)^{-1/m} k \nu_0, \sigma(k)^{1/nm} k^{-1/n} \nu_1, \cdots, \sigma(k)^{1/nm} k^{-1/n} \nu_n)$$

where $(\nu_0, \cdots, \nu_n)$ is in the lattice $(c_0, \cdots, c_n)\Lambda = \Lambda'$. Thus $M(k)$ is the number of latticepoints $(\nu_0, \cdots, \nu_n)$ of Λ' having

$$k/f < \nu_0 \sigma(k)^{-1/m} k \leqq k,$$

$$0 < \nu_0 \sigma(k)^{-1/m} k (L_i(\sigma(k)^{1/nm} k^{-1/n} \nu_1, \cdots, \sigma(k)^{1/nm} k^{-1/n} \nu_n))^n$$
$$\leqq \sigma(\nu_0 \sigma(k)^{-1/m} k) \qquad (1 \leqq i \leqq n)$$

or, what is the same,

$$\sigma(k)^{1/m} f^{-1} < \nu_0 \leqq \sigma(k)^{1/m},$$

$$0 < \nu_0 (L_i(\nu_1, \cdots, \nu_n))^n \leqq \sigma(\nu_0 \sigma(k)^{-1/m} k) \qquad (1 \leqq i \leqq n).$$

Put $y_0 = x_0$ and $y_i = L_i(x_1, \cdots, x_n)$ $(1 \leqq i \leqq n)$. Define the linear transformation τ by $\tau(x_0, \cdots, x_n) = (y_0, \cdots, y_n)$. $M(k)$ is the number of lattice-points of $\tau(\Lambda')$ in the domain $\mathfrak{D}$ of points $(y_0, \cdots, y_n)$ satisfying

$$\sigma(k)^{1/m} f^{-1} < y_0 \leqq \sigma(k)^{1/m}, \tag{3.5}$$

$$0 < y_0 y_i^n \leqq \sigma(\sigma(k)^{-1/m} k y_0) \qquad (1 \leqq i \leqq n). \tag{3.6}$$

The main term for $M(k)$ is the volume $V(\mathfrak{D})$ of $\mathfrak{D}$ divided by the determinant $d(\tau(\Lambda'))$ of $\tau(\Lambda')$. But $d(\tau(\Lambda')) = \mathrm{ad}(\Lambda') = \mathrm{ad}(\Lambda) = a\Delta(R)$, and an easy computation gives

$$V(\mathfrak{D}) = \int_{k/f}^{k} x^{-1} \sigma(x)\, dx.$$

It remains to show

$$|M(k) - V(\mathfrak{D})/d(\tau(\Lambda'))| \leqq C_1 \int_{k/f^2}^{k/f} x^{-1} \sigma(x) \sigma(bx)^{-1/m} dx. \tag{3.7}$$

This we shall do in the next section.

4. Estimate of the error-term. Given a lattice Λ put $\delta(\Lambda)$ for the minimum of the diameters of all fundamental parallelopipeds of Λ.

LEMMA 4. *Let Λ be a lattice, $\mathfrak{D}$ a bounded set in Euclidean E_m, $\mathfrak{D}(\delta)$ where $\delta=\delta(\Lambda)$ the set of points with distance at most δ from the boundary of $\mathfrak{D}$. Assume $\mathfrak{D}$ has Lebesgue measure $V(\mathfrak{D})$ and $\mathfrak{D}(\delta)$ Lebesgue measure $V(\mathfrak{D}(\delta))$. Then the number N of latticepoints of Λ in $\mathfrak{D}$ satisfies*

$$|N-V(\mathfrak{D})/d(\Lambda)|\leqq V(\mathfrak{D}(\delta))/d(\Lambda).$$

Proof. The proof can be left to the reader.

To verify (3.7), we have to estimate $V(\mathfrak{D}(\delta))$ where $\delta=\delta(\tau(\Lambda'))$, $\Lambda'=(c_0,\cdots,c_n)\Lambda$, and where $\mathfrak{D}$ is given by (3.5), (3.6). Since τ and Λ are fixed, and since $|c_i|\leqq C_2{}^n$ $(0\leqq i\leqq n)$, $\delta\leqq\delta_0$.

Since ρ is monotone, we have $\sigma(x')=x'\rho^n(x')\leqq x'\rho^n(x)=(x'/x)\sigma(x)$ for $x'\geqq x$. In particular for $x\leqq x'\leqq fx$, $\sigma(x')\leqq f\sigma(x)$. Thus

$$\sigma(k)^{1/m}=\sigma(k)^{1/n-1/nm}\leqq f^{1/n}\sigma(k/f)^{1/n}\sigma(k)^{-1/nm}=\pi(k),$$

say.

By (3.5), (3.6) the points $(y_0,\cdots,y_n)\in\mathfrak{D}$ have

$$\begin{aligned}0<y_i&\leqq y_0{}^{-1/n}\sigma(\sigma(k)^{-1/m}ky_0)^{1/n}\\&=y_0{}^{-1/n}\sigma(k)^{-1/nm}k^{1/n}y_0{}^{1/n}\rho(\sigma(k)^{-1/m}ky_0)\\&\leqq\sigma(k)^{-1/nm}k^{1/n}\rho(k/f)=f^{1/n}\sigma(k/f)^{1/n}\sigma(k)^{-1/nm}=\pi(k).\\&\qquad\qquad(1\leqq i\leqq n).\end{aligned}$$

Thus

$$0<y_i\leqq\pi(k)\qquad\qquad(0\leqq i\leqq n).$$

The boundary of $\mathfrak{D}$ consists of the subsets $\mathfrak{B}_1$ where $y_0=\sigma(k)^{1/m}f^{-1}$, $\mathfrak{B}_2$ where $y_0=\sigma(k)^{1/m}$, $\mathfrak{B}_{3i}$ where $y_i=0$ $(1\leqq i\leqq n)$ and $\mathfrak{B}_{4i}$ where $y_i{}^ny_0=\sigma(\sigma(k)^{-1/m}ky_0)$ $(1\leqq i\leqq n)$. The points of $\mathfrak{B}_1(\delta_0)$ satisfy

$$|y_0-\sigma(k)^{1/m}f^{-1}|\leqq\delta_0,\ -\delta_0\leqq y_i\leqq\pi(k)+\delta_0\qquad\qquad(1\leqq i\leqq n).$$

Hence $\mathfrak{B}_1(\delta_0)$ has measure at most

$$2\delta_0(\pi(k)+2\delta_0)^n\leqq C_3\pi(k)^n$$

for large k, since $\pi(k)>\sigma(k/f)^{1/n}\sigma(k)^{-1/nm}\geqq f^{-1/nm}\sigma(k/f)^{1/m}>1$ for large k. The sets $\mathfrak{B}_2$, $\mathfrak{B}_{3i}$ are treated similarly. It will suffice to discuss $\mathfrak{B}_{41}$. The points of $\mathfrak{B}_{41}(\delta_0)$ satisfy $-\delta_0\leqq y_i\leqq\pi(k)+\delta_0$ $(2\leqq i\leqq n)$ and (y_0,y_1) is a point in the plane with distance at most δ_0 from the curve $\mathfrak{C}$ given by

$$\sigma(k)^{1/m}f^{-1} \leqq y_0 \leqq \sigma(k)^{1/m},$$
$$y_1 = y_0^{-1/n}\sigma(\sigma(k)^{-1/m}ky_0)^{1/n} = \sigma(k)^{-1/nm}k^{1/n}\rho(\sigma(k)^{-1/m}ky_0).$$

On this curve y_1 is a decreasing function of y_0, and one has $0 \leqq y_j \leqq \pi(k)$ $(j = 0, 1)$. Thus $\mathfrak{C}(\delta_0)$ has (2-dimensional) measure not exceeding $C_4\pi(k)$, and $\mathfrak{B}_{41}(\delta_0)$ (m-dimensional) measure at most $C_5\pi(k)^n$ for large k.

Thus we have shown for $k \geqq k_0$

$$\begin{aligned} V(\mathfrak{D}(\delta)) &\leqq C_6\pi(k)^n \\ &= C_6 f\sigma(k/f)\sigma(k)^{-1/m} \\ &\leqq C_7\int_{k/f^2}^{k/f} x^{-1}\sigma(x)\sigma(f^2x)^{-1/m}dx, \end{aligned}$$

and (3.7) follows.

5. Proof of Theorem 1. Let $\beta_1, \cdots, \beta_n$ be as in Theorem 1, F the numberfield with basis $\beta_0 = 1, \beta_1, \cdots, \beta_n$. Choose a basis $\alpha_0, \cdots, \alpha_n$ of F according to Lemma 2. Let R be the module with basis $\alpha_0, \cdots, \alpha_n$. With $\beta_0 = 1$ (2.5) implies

$$\alpha_0^{(i)} + \alpha_1^{(i)}\beta_1 + \cdots + \alpha_n^{(i)}\beta_n = 0 \qquad (1 \leqq i \leqq n),$$

whence

$$\alpha_0^{*(i)} + \alpha_0^{*(i)}\beta_1 + \cdots + \alpha_n^{*(i)}\beta_n = 0 \qquad (1 \leqq i \leqq n). \tag{5.1}$$

With integers $q; p_1, \cdots, p_n$ we associate

$$\xi = -\alpha_0 q + \alpha_1 p_1 + \cdots + \alpha_n p_n \in R.$$

$$\xi^{*(i)} = -\alpha_0^{*(i)}q + \alpha_1^{*(i)}p_1 + \cdots + \alpha_n^{*(i)}p_n \qquad (0 \leqq i \leqq n). \tag{5.2}$$

Combining this equation with (5.1) we obtain

$$\xi^{*(i)} = \alpha_1^{*(i)}(\beta_1 q + p_1) + \cdots + \alpha_n^{*(i)}(\beta_n q + p_n) \qquad (1 \leqq i \leqq n)$$

The nonvanishing of the determinant δ of Lemma 2 implies the nonvanishing of the determinant $\delta^* = |\alpha_j^{*(i)}|$ $(1 \leqq i, j \leqq n)$. Denote the inverse matrix to $(\alpha_j^{*(i)})$ $(1 \leqq i, j \leqq n)$ by (c_{ij}) $(1 \leqq i, j \leqq n)$. Define linear forms L_i by

$$L_i(x_1, \cdots, x_n) = \sum_{j=1}^{n} c_{ij}x_j \qquad (1 \leqq i \leqq n).$$

Then

$$\beta_i q + p_i = \sum_{j=1}^{n} c_{ij}\xi^{*(j)} = L_i(\xi^{*(1)}, \cdots, \xi^{*(n)}) \qquad (1 \leqq i \leqq n).$$

Therefore

$$0 < \beta_i q + p_i \leqq \psi(q) \qquad (1 \leqq i \leqq n) \tag{5.3}$$

holds if and only if

$$0 < L_i(\xi^{*(1)}, \cdots, \xi^{*(n)}) \leqq \psi(q) \qquad (1 \leqq i \leqq n) \tag{5.4}$$

holds.

Solving (5.2) for q we find

$$q = d_0 \xi^{*(0)} + \cdots + d_n \xi^{*(n)} \tag{5.5}$$

where $d_0 = \pm \delta^* \Delta(R)^{-1}$. If $d_0 > 0$ we are lucky. If not, simply replace $\alpha_0, \cdots, \alpha_n$ by $-\alpha_0, \cdots, -\alpha_n$, do everything else as before, and d_0 will become positive. Hence assume $d_0 > 0$. Now (5.3) or (5.4) implies $0 < L_i(\xi^{*(1)}, \cdots, \xi^{*(n)}) \leqq \psi(1)$ $(1 \leqq i \leqq n)$, hence $\xi^{*(1)}, \cdots, \xi^{*(n)}$ are bounded, and by (5.5)

$$| q - d_0 \xi^{*(0)} | \leqq C_8,$$

Extend $\psi(q)$ to a montone function defined for all reals x and with $\psi(x) = \psi(1)$ for $x \leqq 1$. Put $\omega(x) = x\psi^n(x)$,

$$\psi_j(x) = \psi(d_0 x - (-1)^j C_8) \qquad (j = 1, 2).$$

One has $\psi_1(x) \leqq \psi(x) \leqq \psi_2(x)$ and

$$N_1(h) \leqq N(h; \beta_1, \cdots, \beta_n) \leqq N_2(h),$$

where $N_j(h)$ is the number of $\xi \in R$ with

$$\begin{aligned} &1 - (-1)^j C_8/d_0 \leqq \xi^{*(0)} \leqq h/d_0 + (-1)^j C_8/d_0, \\ &0 < L_i(\xi^{*(1)}, \cdots, \xi^{*(n)}) \leqq \psi_j(\xi^{*(0)}) \qquad (1 \leqq i \leqq n). \end{aligned} \tag{5.7}$$

The functions $\rho_j(x) = \psi_j(x)$ decrease, while $\sigma_j(x) = x\rho_j(x)^n$ tend to infinity $(j = 1, 2)$. Put $b = \sqrt{2}$. Observing $a^{-1}\Delta(R)^{-1} = | \delta^* | \Delta(R)^{-1} = d_0$ and applying Lemma 3 one obtains

$$N_j(h) = d_0 \int_1^{h/d_0} x^{-1}\sigma_j(x)\, dx + 0\Big(\int_1^{h/d_0} x^{-1}\sigma_j(x)\sigma_j(\sqrt{2}x)^{-1/m} dx\Big). \tag{5.8}$$

(This formula holds for $0 < \xi^{*(0)} \leqq h/d_0$ instead of (5.7), as shown by Lemma 3. But it is easy to see that our taking (5.7) rather than $0 < \xi^{*(0)} \leqq h/d_0$ involves an error $0(1)$ which is absorbed by the 0 of (5.8)).

$$\begin{aligned} d_0 \int_1^{h/d_0} x^{-1}\sigma_j(x)\, dx &= d_0 \int_1^{h/d_0} \psi_j(x)^n dx = d_0 \int_1^{h/d_0} \psi(d_0 x)^n dx + 0(1) \\ &= \int_1^h \psi(x)^n dx + 0(1) = \Psi(h) + 0(1). \end{aligned}$$

As for the error-term in (5.8),

$$\int_1^{h/d_0} x^{-1}\sigma_j(x)\sigma_j(\sqrt{2}x)^{-1/m}dx$$

$$= \int_1^{h/d_0} \rho_j^n(x)(\sqrt{2}x)^{-1/m}\rho_j(\sqrt{2}x)^{-n/m}dx$$

$$= 0\Big(\int_1^{h/d_0} \psi(d_0x)^n x^{-1/m}\psi(\sqrt{2}d_0x)^{-n/m}dx\Big)$$

$$= 0\Big(\int_1^{h} \psi(x)^n\omega(\sqrt{2}x)^{-1/m}dx\Big)$$

$$= 0\Big(\sum_{q=1}^{h} \psi(q)^n\omega(2q)^{-1/m}dx\Big).$$

This proves the theorem.

UNIVERSITY OF COLORADO,
BOULDER, COLORADO.

REFERENCES.

[1] P. Erdös, "Some results on diophantine approximation," *Acta Arithmetica,* vol. 5 (1959), pp. 359-369.

[2] S. Lang, "Asymptotic approximation to quadratic irrationalities," *American Journal of Mathematics,* vol. 87 (1965), pp. 481-487.

[3] ———, "Asymptotic Approximation to quadratic irrationalities II," *American Journal of Mathematics,* vol. 87 (1965), pp. 488-496.

[4] W. M. Schmidt, "A metrical theorem in diophantine approximation," *Canadian Journal of Mathematics,* vol. 12 (1960), pp. 619-631.

CATEGORIES OF FUNCTORS AND ADJOINT FUNCTORS.

By M. André.[*]

If $\mathfrak{A}$ is a small category (the class of all its morphisms is a set) and if $\mathfrak{M}$ is a category, we may consider the category $(\mathfrak{A},\mathfrak{M})$: the objects are all covariant functors from $\mathfrak{A}$ to $\mathfrak{M}$, the morphisms are all natural transformations of such functors and the composition law is that of such transformations. The category $(\mathfrak{A},\mathfrak{M})$ is a *category of functors.* In fact, when we say "the category of functors $(\mathfrak{A},\mathfrak{M})$," we do not think of the category alone; we think of the category described by means of the categories $\mathfrak{A}$ and $\mathfrak{M}$. In a certain sense, a category of functors is a presentation of a category. The main purpose of the theory of categories of functors is the reduction of problems concerning one category $(\mathfrak{A},\mathfrak{M})$ or two categories $(\mathfrak{A},\mathfrak{M})$ and $(\mathfrak{B},\mathfrak{N})$ to problems concerning the category $\mathfrak{M}$ or the categories $\mathfrak{M}$ and $\mathfrak{N}$.

Two categories $\mathfrak{P}$ and $\mathfrak{Q}$ being given, two functors $T\colon \mathfrak{P}\rightarrow\mathfrak{Q}$ and $S\colon \mathfrak{Q}\rightarrow\mathfrak{P}$ are *adjoint,* T on the left and S on the right, if and only if the following functors are equivalent (contravariant in $\mathfrak{P}$, covariant in $\mathfrak{Q}$ with values in the category of all sets): to the object P of $\mathfrak{P}$ and to the object Q of $\mathfrak{Q}$ there correspond the sets $\mathrm{Hom}_{\mathfrak{Q}}(TP,Q)$ and $\mathrm{Hom}_{\mathfrak{P}}(P,SQ)$. See [1] for further details. The main purpose of this paper is the study of the adjoint functors between $(\mathfrak{A},\mathfrak{M})$ and $(\mathfrak{B},\mathfrak{N})$ by means of the adjoint functors between $\mathfrak{M}$ and $\mathfrak{N}$. We want to answer the two following connected questions. How to construct the pairs of adjoint functors between $(\mathfrak{A},\mathfrak{M})$ and $(\mathfrak{B},\mathfrak{N})$ by means of the pairs of adjoint functors between $\mathfrak{M}$ and $\mathfrak{N}$? How to decide whether a functor from $(\mathfrak{A},\mathfrak{M})$ to $(\mathfrak{B},\mathfrak{N})$ has a right or left adjoint by means of the solution of the same problem concerning functors from $\mathfrak{M}$ to $\mathfrak{N}$?

The general procedure is the following. We replace a functor from $(\mathfrak{A},\mathfrak{M})$ to $(\mathfrak{B},\mathfrak{N})$ by a set of functors from $\mathfrak{M}$ to $\mathfrak{N}$, and the original functor has an adjoint essentially if and only if each functor of the set has an adjoint. This procedure allows us to study the following explicit problems, as examples. If a functor from $\mathfrak{A}$ to $\mathfrak{B}$ is given, has the corresponding functor from $(\mathfrak{B},\mathfrak{M})$ to $(\mathfrak{A},\mathfrak{M})$ a right or left adjoint? If a small category $\mathfrak{A}$ and a category $\mathfrak{M}$ are given besides the category $\mathfrak{E}$ of all sets and maps, when does a functor from $(\mathfrak{A},\mathfrak{E})$ to $\mathfrak{M}$ have a right adjoint (see [2])? The two functors

---

Received July 31, 1964.

[*] This research was partly supported by the U. S. Department of Army through its European Research Office.

appearing in the Dold-Kan-Puppe equivalence (see [3]) are adjoint; can they be considered as a special case of a general construction?

This paper is divided into three parts, the first two dealing with the general theory and the last part with examples:

I. Categories of functors and complete functors.
II. Categories of functors and adjoint functors.
III. Examples of adjoint functors.

A short summary of this paper might be the following. *The property of having an adjoint for a functor from any category to a category of functors is a "local" property.*

I. Categories of Functors and Complete Functors.

It is well known that a left adjoint functor preserves right roots (or limits). Consequently, a first step in the study of adjoint functors between categories of functors is that of root preserving functors between categories of functors. The scheme of the theory is the following. To each functor from $(\mathfrak{A},\mathfrak{M})$ to $\mathfrak{N}$ there corresponds a functor from $\mathfrak{M}$ to $(\mathfrak{A}^{dual},\mathfrak{N})$ and to each functor from $\mathfrak{M}$ to $(\mathfrak{A}^{d},\mathfrak{N})$ there corresponds a functor from $(\mathfrak{A},\mathfrak{M})$ to $\mathfrak{N}$. These two constructions—right localisation and right globalisation—are inverse to each other in case of right root preserving functors, and a functor from $\mathfrak{M}$ to $(\mathfrak{A}^{d},\mathfrak{N})$ is more easily dealt with than a functor from $(\mathfrak{A},\mathfrak{M})$ to $\mathfrak{N}$: consequently, we gain something in the reduction of problems concerning categories of functors $(\mathfrak{A},\mathfrak{M})$, etc., to problems concerning the categories $\mathfrak{M}$, etc.

1. Right complete functors. Let $\mathfrak{M}$ be a category and $\mathfrak{A}$ a small category. Let us consider the category $(\mathfrak{A},\mathfrak{M})$ of all covariant functors from $\mathfrak{A}$ to $\mathfrak{M}$ and the functor $L: \mathfrak{M} \to (\mathfrak{A},\mathfrak{M})$ corresponding to the projection functor of the product category $\mathfrak{M} \times \mathfrak{A}$ onto $\mathfrak{M}$. In other words, for any object M of $\mathfrak{M}$, the functor $L(M)$ lets the identity morphism of M correspond to any morphism of $\mathfrak{A}$. A first simple question concerning adjoint functors of categories of functors appears: has the functor L a left adjoint? If the left adjoint functor of L exists, this functor is quoted as $^*L$, and its existence is an important property of the category $\mathfrak{M}$. If for any small category $\mathfrak{A}$, the functor L has a left adjoint, then the category $\mathfrak{M}$ is called *right complete*: see [4]. For instance, if the small category $\mathfrak{A}$ consists only of identities, the functor $^*L$ is a direct sum functor.

Furthermore, a functor $F: \mathfrak{M} \to \mathfrak{N}$ between two right complete categories is called *right complete* if and only if the following diagram of functors is commutative for any small category $\mathfrak{A}$ (right root preserving functor):

$$\begin{tikzcd}
(\mathfrak{A}, \mathfrak{M}) \arrow[r, "*L"] \arrow[d, "(\mathfrak{A}, F)"] & \mathfrak{M} \arrow[d, "F"] \\
(\mathfrak{A}, \mathfrak{N}) \arrow[r, "*L"] & \mathfrak{N}
\end{tikzcd}$$

There are the dual notions and properties: functor $L^*: (\mathfrak{A}, \mathfrak{M}) \to \mathfrak{M}$, left complete category, direct product and left complete functor.

2. Right localisation. Let $\mathfrak{M}$ and $\mathfrak{N}$ be two right complete categories, and $\mathfrak{A}$ a small category. Let us consider a functor F from $(\mathfrak{A}, \mathfrak{M})$ to $\mathfrak{N}$. The *right localisation* of F is a functor $R\operatorname{loc}(F)$ from $\mathfrak{M}$ to $(\mathfrak{A}^d, \mathfrak{N})$ defined as follows. The category $\mathfrak{A}^d$ is the dual category of the small category $\mathfrak{A}$.

First of all, let us consider the category $\mathfrak{E}$ of all sets and maps. There are the functor $\operatorname{Hom}_{\mathfrak{A}}$ from the product category $\mathfrak{A}^d \times \mathfrak{A}$ to the category $\mathfrak{E}$, and the functor $\sum$ from the product category $\mathfrak{M} \times \mathfrak{E}$ to the category $\mathfrak{M}$: to an object $M \times E$ there corresponds the object $\sum_E M$; to a morphism $m \times E$ there corresponds the morphism $\sum_E m$, and to a morphism $M \times e$ there corresponds the morphism whose composition with the injection morphism of M into $\sum_E M$ corresponding to an element ϵ of E is equal to the injection morphism of M into $\sum_{E'} M$ corresponding to the element $e(\epsilon)$ of E'. By composition, we get a functor from $\mathfrak{M} \times \mathfrak{A}^d \times \mathfrak{A}$ to $\mathfrak{M}$ equal to

$$\sum \circ (\mathrm{Id}_{\mathfrak{M}} \times \operatorname{Hom}_{\mathfrak{A}}): \mathfrak{M} \times \mathfrak{A}^d \times \mathfrak{A} \to \mathfrak{M} \times \mathfrak{E} \to \mathfrak{M}.$$

This functor corresponds to a functor $*J$ from $\mathfrak{M}$ to $(\mathfrak{A}^d \times \mathfrak{A}, \mathfrak{M})$.

By definition, the functor $R\operatorname{loc}(F)$ is equal to the following composition:

$$\mathfrak{M} \xrightarrow{*J} (\mathfrak{A}^d \times \mathfrak{A}, \mathfrak{M}) \cong (\mathfrak{A}^d, (\mathfrak{A}, \mathfrak{M})) \xrightarrow{(\mathfrak{A}^d, F)} (\mathfrak{A}^d, \mathfrak{N}).$$

The functor $*J$ being defined by means of direct sums is right complete. Furthermore, the functor $(\mathfrak{A}^d, F)$ is right complete if the functor F is so. Consequently:

2.1. *The right localisation of a right complete functor is a right complete functor.*

3. Right globalisation. Let $\mathfrak{M}$ and $\mathfrak{N}$ be two right complete categories, and $\mathfrak{A}$ a small category. Let us consider a functor G from $\mathfrak{M}$ to $(\mathfrak{A}^d, \mathfrak{N})$. The *right globalisation* of G is a functor $R\,\mathrm{glo}(G)$ from $(\mathfrak{A}, \mathfrak{M})$ to $\mathfrak{N}$ defined as follows.

First of all, let us consider the small category $\mathfrak{A}$ alone. Besides the product category $\mathfrak{A} \times \mathfrak{A}^d$, there is another interesting small category $\hat{\mathfrak{A}}$ defined as follows. The objects of $\hat{\mathfrak{A}}$ are all the morphisms of $\mathfrak{A}$, the morphisms of $\mathrm{Hom}_{\hat{\mathfrak{A}}}(\alpha, \beta)$ are all the pairs $(a, \bar{a})$ of morphisms of $\mathfrak{A}$ such that $\bar{a}\beta a$ is defined and equal to α and the composition law of $\hat{\mathfrak{A}}$ maps $(a, \bar{a})$ of $\mathrm{Hom}_{\hat{\mathfrak{A}}}(\alpha, \beta)$ and $(c, \bar{c})$ of $\mathrm{Hom}_{\hat{\mathfrak{A}}}(\beta, \gamma)$ onto $(ca, \bar{a}\bar{c})$ of $\mathrm{Hom}_{\hat{\mathfrak{A}}}(\alpha, \gamma)$:

$$\begin{tikzcd}
A \arrow[r, "a"] \arrow[d, "\alpha"'] & B \arrow[r, "c"] \arrow[d, "\beta"'] & C \arrow[d, "\gamma"'] \\
\bar{A} & \bar{B} \arrow[l, "\bar{a}"'] & \bar{C} \arrow[l, "\bar{c}"']
\end{tikzcd}$$

There is of course a functor K from $\hat{\mathfrak{A}}$ to $\mathfrak{A} \times \mathfrak{A}^d$: the image $K(a, \bar{a})$ of the pair of morphisms a and $\bar{a}$ of $\mathfrak{A}$ considered as a morphism of $\hat{\mathfrak{A}}$ is equal to the same pair but considered as a morphism of $\mathfrak{A} \times \mathfrak{A}^d$.

By definition the functor $R\,\mathrm{glo}(G)$ is equal to the following composition:

$$(\mathfrak{A}, \mathfrak{M}) \xrightarrow{(\mathfrak{A}, G)} (\mathfrak{A}, (\mathfrak{A}^d, \mathfrak{N})) \cong (\mathfrak{A} \times \mathfrak{A}^d, \mathfrak{N}) \xrightarrow{(K, \mathfrak{N})} (\hat{\mathfrak{A}}, \mathfrak{N}) \xrightarrow{{}^*L} \mathfrak{N}$$

The functors $(K, \mathfrak{N})$ and ${}^*L$ are right complete, of course. Furthermore, the functor $(\mathfrak{A}, G)$ is right complete if the functor G is so. Consequently:

3.1. *The right globalisation of a right complete functor is a right complete functor.*

4. Right localisation and globalisation are inverse. Let $\mathfrak{M}$ and $\mathfrak{N}$ be two right complete categories and $\mathfrak{A}$ a small category. We shall prove that for right complete functors from $(\mathfrak{A}, \mathfrak{M})$ to $\mathfrak{N}$ or from $\mathfrak{M}$ to $(\mathfrak{A}^d, \mathfrak{N})$ the right localisation and the right globalisation are inverse constructions: see 2.1 and 3.1. Let us begin with a lemma.

LEMMA 4.1. *The following composition of functors is equivalent to the identity functor of* $(\mathfrak{A}, \mathfrak{M})$:

$$\begin{tikzcd}
(\mathfrak{A}, \mathfrak{M}) \arrow[r, "(\mathfrak{A}, {}^*J)"] & (\mathfrak{A}, (\mathfrak{A}^d \times \mathfrak{A}, \mathfrak{M})) \cong (\mathfrak{A} \times \mathfrak{A}^d, (\mathfrak{A}, \mathfrak{M})) \arrow[d, "{(K, (\mathfrak{A}, \mathfrak{M}))}"] \\
(\mathfrak{A}, \mathfrak{M}) & (\hat{\mathfrak{A}}, (\mathfrak{A}, \mathfrak{M})) \arrow[l, "{}^*L_{(\mathfrak{A}, \mathfrak{M})}"']
\end{tikzcd}$$

Remark. The order of the two exemplars of the category $\mathfrak{A}$ is *not* inverted in the equivalence:

$$(\mathfrak{A}, (\mathfrak{A}^d \times \mathfrak{A}, \mathfrak{M})) \cong (\mathfrak{A} \times \mathfrak{A}^d, (\mathfrak{A}, \mathfrak{M}))$$

Proof. Let us consider a functor F from $\mathfrak{A}$ to $\mathfrak{M}$. Its image is another functor $\bar{F}$ from $\mathfrak{A}$ to $\mathfrak{M}$ defined as follows. The object $\bar{F}(A)$ corresponding to an object A of $\mathfrak{A}$ is the right root of the following functor from $\hat{\mathfrak{A}}$ to $\mathfrak{M}$: to the object $\omega\colon X \rightarrow Y$ of $\hat{\mathfrak{A}}$ there corresponds the object

$$\sum_{\mathrm{Hom}(Y,A)} F(X)$$

covariant in X and contravariant in Y, consequently natural with respect to $\hat{\mathfrak{A}}$. There is a natural morphism, for any $\omega\colon X \rightarrow Y$,

$$\sum_{\mathrm{Hom}(Y,A)} F(X) \rightarrow F(A)$$

this morphism, composed with the injection morphism

$$F(X) \rightarrow \sum_{\mathrm{Hom}(Y,A)} F(X)$$

corresponding to the morphism η of $\mathrm{Hom}(Y, A)$, is equal to the morphism $F(\eta\omega)$. By means of this morphism it is easy to prove that $F(A)$ is the right root of the functor above. Consequently, the objects $F(A)$ and $\bar{F}(A)$ are equivalent in a natural way. In other words, the functors F and $\bar{F}$ are equivalent, and the lemma is proved.

The following result is of the same type, and can be deduced from Lemma 4.1.

LEMMA 4.2. *The following composition of functors is equivalent to the identity functor of* $(\mathfrak{A}^d, \mathfrak{N})$:

$$\begin{tikzcd}
(\mathfrak{A}^d, \mathfrak{N}) \arrow[r, "{*J_{(\mathfrak{A}^d,\mathfrak{N})}}"] & (\mathfrak{A}^d \times \mathfrak{A}, (\mathfrak{A}^d, \mathfrak{N})) \arrow[r, phantom, "\cong"] & (\mathfrak{A}^d, (\mathfrak{A} \times \mathfrak{A}^d, \mathfrak{N})) \arrow[d, "{(\mathfrak{A}^d, (K, \mathfrak{N}))}"] \\
 & (\mathfrak{A}^d, \mathfrak{N}) & (\mathfrak{A}^d, (\hat{\mathfrak{A}}, \mathfrak{N})) \arrow[l, "{(\mathfrak{A}^d, *L)}"']
\end{tikzcd}$$

Proof. The composition of functors above is equal to the following one:

$$\begin{tikzcd}
(\mathfrak{A}^d, \mathfrak{N}) \arrow[r, "{(\mathfrak{A}^d, *J)}"] & (\mathfrak{A}^d, (\mathfrak{A}^d \times \mathfrak{A}, \mathfrak{N})) \arrow[r, phantom, "\cong"] & (\mathfrak{A} \times \mathfrak{A}^d, (\mathfrak{A}^d, \mathfrak{N})) \arrow[d, "{(K, (\mathfrak{A}^d, \mathfrak{N}))}"] \\
 & (\mathfrak{A}^d, \mathfrak{N}) & (\hat{\mathfrak{A}}, (\mathfrak{A}^d, \mathfrak{N})) \arrow[l, "{*L_{(\mathfrak{A}^d,\mathfrak{N})}}"']
\end{tikzcd}$$

This is exactly the composition of functors appearing in Lemma 4.1 rewritten for contravariant functors.

And now we are able to prove two theorems in which it appears that the right localisation and the right globalisation are inverse constructions.

THEOREM 4.3. *Let* $\mathfrak{M}$ *and* $\mathfrak{N}$ *be two right complete categories,* $\mathfrak{A}$ *a small category and* F *a right complete functor from* $(\mathfrak{A}, \mathfrak{M})$ *to* $\mathfrak{N}$*. Then the functor* $R\operatorname{loc}(F)$ *from* $\mathfrak{M}$ *to* $(\mathfrak{A}^d, \mathfrak{N})$ *is right complete and the functor* $R\operatorname{glo}[R\operatorname{loc}(F)]$ *is equivalent to* F*.*

Proof. The first part of the theorem is obvious: see 2.1.

For the proof of the second part of the theorem we have to compute the functor $R\operatorname{glo}[R\operatorname{loc}(F)]$ given by the following composition of functors:

$$\begin{tikzcd}
(\mathfrak{A}, \mathfrak{M}) \arrow[r, "{(\mathfrak{A}, {}^*J)}"] & (\mathfrak{A}, (\mathfrak{A}^d \times \mathfrak{A}, \mathfrak{M})) \arrow[r, phantom, "\cong"] & (\mathfrak{A} \times \mathfrak{A}^d, (\mathfrak{A}, \mathfrak{M})) \arrow[d, "{(\mathfrak{A} \times \mathfrak{A}^d, F)}"] \\
\mathfrak{N} & (\hat{\mathfrak{A}}, \mathfrak{N})) \arrow[l, "{}^*L"'] & (\mathfrak{A} \times \mathfrak{A}^d, \mathfrak{N}) \arrow[l, "{(K, \mathfrak{N})}"']
\end{tikzcd}$$

Since the functor F is right complete, this composition of functors can be replaced by the following one:

$$\begin{tikzcd}
(\mathfrak{A}, \mathfrak{M}) \arrow[r, "{(\mathfrak{A}, {}^*J)}"] & (\mathfrak{A}, (\mathfrak{A}^d \times \mathfrak{A}, \mathfrak{M})) \arrow[r, phantom, "\cong"] & (\mathfrak{A} \times \mathfrak{A}^d, (\mathfrak{A}, \mathfrak{M})) \arrow[d, "{(K, (\mathfrak{A}, \mathfrak{M}))}"] \\
\mathfrak{N} & (\mathfrak{A}, \mathfrak{M}) \arrow[l] & (\hat{\mathfrak{A}}, (\mathfrak{A}, \mathfrak{M})) \arrow[l, "{{}^*L_{(\mathfrak{A}, \mathfrak{M})}}"']
\end{tikzcd}$$

By Lemma 4.1, this functor is equivalent to the functor F.

THEOREM 4.4. *Let* $\mathfrak{M}$ *and* $\mathfrak{N}$ *be two right complete categories,* $\mathfrak{A}$ *a small category and* G *a right complete functor from* $\mathfrak{M}$ *to* $(\mathfrak{A}^d, \mathfrak{N})$*. Then the functor* $R\operatorname{glo}(G)$ *from* $(\mathfrak{A}, \mathfrak{M})$ *to* $\mathfrak{N}$ *is right complete and the functor* $R\operatorname{loc}[R\operatorname{glo}(G)]$ *is equivalent to* G*.*

Proof. The first part of the theorem is obvious: see 2.2.

For the proof of the second part of the theorem we have to compute the functor $R\operatorname{loc}[R\operatorname{glo}(G)]$ given by the following composition of functors:

$$\begin{tikzcd}
\mathfrak{M} \arrow[r, "{}^*J"] & (\mathfrak{A}^d \times \mathfrak{A}, \mathfrak{M}) \arrow[r, "{(\mathfrak{A}^d \times \mathfrak{A}, G)}"] & (\mathfrak{A}^d \times \mathfrak{A}, (\mathfrak{A}^d, \mathfrak{N})) \arrow[d, phantom, "\cong"] \\
(\mathfrak{A}^d, \mathfrak{N}) & (\mathfrak{A}^d, (\hat{\mathfrak{A}}, \mathfrak{N})) \arrow[l, "{(\mathfrak{A}^d, {}^*L)}"'] & (\mathfrak{A}^d, (\mathfrak{A} \times \mathfrak{A}^d, \mathfrak{N})) \arrow[l, "{(\mathfrak{A}^d, (K, \mathfrak{N}))}"']
\end{tikzcd}$$

Since the functor G is right complete, this composition of functors can be replaced by the following one:

$$\begin{tikzcd}
\mathfrak{M} \arrow[r, "G"] & (\mathfrak{A}^d, \mathfrak{N}) \arrow[r, "*J_{(\mathfrak{A}^d, \mathfrak{N})}"] & (\mathfrak{A}^d \times \mathfrak{A}, (\mathfrak{A}^d, \mathfrak{N})) \arrow[d, phantom, "\cong"] \\
(\mathfrak{A}^d, \mathfrak{N}) & (\mathfrak{A}^d, (\hat{\mathfrak{A}}, \mathfrak{N})) \arrow[l, "(\mathfrak{A}^d, *L)"'] & (\mathfrak{A}^d, (\mathfrak{A} \times \mathfrak{A}^d, \mathfrak{N})) \arrow[l, "(\mathfrak{A}^d, (K, \mathfrak{N}))"']
\end{tikzcd}$$

By Lemma 4.2, this functor is equivalent to the functor G, and the theorem is proved.

Thus, the right localisation and the right globalisation are inverse constructions for right complete functors; we are now able to study adjoint functors.

II. Categories of Functors and Adjoint Functors.

By means of the right localisation and the right globalisation we can answer the two complementary questions about adjoint functors. When has a functor from $\mathfrak{N}$ to $(\mathfrak{A}, \mathfrak{M})$ a left adjoint, and when has a functor from $(\mathfrak{A}, \mathfrak{M})$ to $\mathfrak{N}$ a right adjoint? The first functor has a left adjoint if and only if it has a left adjoint in a weaker meaning (local adjoint), and the second functor has a right adjoint if and only if it is right complete and has a right localisation with a right adjoint in a weaker meaning (local adjoint). The problem of the existence of a left or right local adjoint from $\mathfrak{M}$ to $(\mathfrak{A}, \mathfrak{N})$ is equivalent to a problem of the existence of left or right adjoints for functors from $\mathfrak{M}$ to $\mathfrak{N}$; consequently, the aim of the theory is reached.

5. Section functors and local adjoints. Let $\mathfrak{M}$ be a right complete category and $\mathfrak{A}$ a small category. For any object A of $\mathfrak{A}$ there is a simple functor Γ_A from $(\mathfrak{A}, \mathfrak{M})$ to $\mathfrak{M}$: *the section functor* Γ_A maps the functor F of $(\mathfrak{A}, \mathfrak{M})$ onto the object $F(A)$ of $\mathfrak{M}$. The section functors have left adjoints.

LEMMA 5.1. *The section functor* $\Gamma_A : (\mathfrak{A}, \mathfrak{M}) \rightarrow \mathfrak{M}$ *has a left adjoint.*

Proof. This result is well known. Actually the left adjoint functor of Γ_A is equal to the following composition of functors:

$$\mathfrak{M} \xrightarrow{*J} (\mathfrak{A}^d \times \mathfrak{A}, \mathfrak{M}) \cong (\mathfrak{A}^d, (\mathfrak{A}, \mathfrak{M})) \xrightarrow{\Gamma_A} (\mathfrak{A}, \mathfrak{M}).$$

Here we use the section functor Γ_A for the small category $\mathfrak{A}^d$ and the category $(\mathfrak{A}, \mathfrak{M})$.

Let $\mathfrak{M}$ and $\mathfrak{N}$ be two categories and $\mathfrak{A}$ a small category. Two functors

G from $\mathfrak{M}$ to $(\mathfrak{A}^d, \mathfrak{N})$ and H from $\mathfrak{N}$ to $(\mathfrak{A}, \mathfrak{M})$ are *locally adjoint* if and only if for any object A of $\mathfrak{A}$ the functors $\Gamma_A \circ G$ from $\mathfrak{M}$ to $\mathfrak{N}$ and $\Gamma_A \circ H$ from $\mathfrak{N}$ to $\mathfrak{M}$ are adjoint. It is clear that a functor G from $\mathfrak{M}$ to $(\mathfrak{A}^d, \mathfrak{N})$ has a right local adjoint if and only if for any object A of $\mathfrak{A}$ the functor $\Gamma_A \circ G$ has a right adjoint and that this local right adjoint $lR\,\mathrm{ad}(G)$ is determined by G up to an equivalence.

The key of the theory is the following proposition.

PROPOSITION 5.2. *Let $\mathfrak{M}$ and $\mathfrak{N}$ be two right complete categories and $\mathfrak{A}$ a small category. Let G be a functor from $\mathfrak{M}$ to $(\mathfrak{A}^d, \mathfrak{N})$ with a local right adjoint. Then the functors $R\,\mathrm{glo}(G)$ from $(\mathfrak{A}, \mathfrak{M})$ to $\mathfrak{N}$ and $lR\,\mathrm{ad}(G)$ from $\mathfrak{N}$ to $(\mathfrak{A}, \mathfrak{M})$ are left and right adjoint functors.*

Proof. Let X be an object of $(\mathfrak{A}, \mathfrak{M})$ and Y an object of $\mathfrak{N}$. We have to consider the two following sets:

$$\mathrm{Hom}_{\mathfrak{N}}[R\,\mathrm{glo}(G)(X), Y] \quad \text{and} \quad \mathrm{Hom}_{(\mathfrak{A},\mathfrak{M})}[X, lR\,\mathrm{ad}(G)(Y)].$$

Both these sets may be computed by means of the composition of functors:

$$(\mathfrak{A}^d \times \mathfrak{A}, \mathfrak{E}) \xrightarrow{(K^d, \mathfrak{E})} (\hat{\mathfrak{A}}^d, \mathfrak{E}) \xrightarrow{L^*} \mathfrak{E}.$$

$\mathfrak{E}$ being the category of all sets and maps, and L^* the functor left root. The definition of the right globalisation shows that the first set is equal to the image of the following functor from $\mathfrak{A}^d \times \mathfrak{A}$ to $\mathfrak{E}$ by the functor $L^* \circ (K^d, \mathfrak{E})$: to A of $\mathfrak{A}$ and $\bar{A}$ of $\mathfrak{A}^d$ there corresponds the set:

$$\mathrm{Hom}_{\mathfrak{N}}(\Gamma_A \circ G \circ \Gamma_{\bar{A}}(X), Y).$$

The other classical method of computing $\mathrm{Hom}_{(\mathfrak{A},\mathfrak{M})}$ (see [1] for instance) shows that the second set above is equal to the image of the following functor from $\mathfrak{A}^d \times \mathfrak{A}$ to $\mathfrak{E}$ by the functor $L^* \circ (K, \mathfrak{E})$: to A of $\mathfrak{A}$ and $\bar{A}$ of $\mathfrak{A}^d$ there corresponds the set:

$$\mathrm{Hom}_{\mathfrak{M}}(\Gamma_{\bar{A}}(X), \Gamma_A \circ lR\,\mathrm{ad}(G)(Y)).$$

Since the functors $\Gamma_A \circ G$ and $\Gamma_A \circ lR\,\mathrm{ad}(G)$ are adjoint functors (see the definition of locally adjoint functors), the two considered functors from $\mathfrak{A}^d \times \mathfrak{A}$ to $\mathfrak{E}$ are equivalent in a natural way. Consequently, the two sets of the beginning of the proof are equivalent in a natural way with respect to X and Y. In other words, the functors $R\,\mathrm{glo}(G)$ and $lR\,\mathrm{ad}(G)$ are adjoint functors. The proposition is proved.

6. Categories of functors and adjoint functors. We are now able to prove the final results which complete Theorems 4.3 and 4.4. Let us start with the easier, which does not use these theorems.

THEOREM 6.1. *Let* $\mathfrak{M}$ *and* $\mathfrak{N}$ *be two right complete categories,* $\mathfrak{A}$ *a small category and* H *a functor from* $\mathfrak{N}$ *to* $(\mathfrak{A}, \mathfrak{M})$. *Then the functor* H *has a left adjoint if and only if it has a local left adjoint; in this case the left adjoint functor of* H *is equal to* $R\,\mathrm{glo}[lL\,\mathrm{ad}(H)]$, *the right globalisation of the local left adjoint of* H.

Proof. The functor Γ_A from $(\mathfrak{A}, \mathfrak{M})$ to $\mathfrak{M}$ has a left adjoint: see Lemma 5.1. Consequently, if the functor H has a left adjoint, the functor $\Gamma_A \circ H$ has a left adjoint too. Thus the functor H has a local left adjoint.

Conversely, if the functor H has a local left adjoint, we may apply Proposition 5.2 to the functor $lL\,\mathrm{ad}(H)$. The functors $R\,\mathrm{glo}[lL\,\mathrm{ad}(H)]$ and $lR\,\mathrm{ad}[lL\,\mathrm{ad}(H)] = H$ are adjoint, and the theorem is completely proved.

And now let us prove the complementary result.

THEOREM 6.2. *Let* $\mathfrak{M}$ *and* $\mathfrak{N}$ *be two right complete categories,* $\mathfrak{A}$ *a small category and* F *a functor from* $(\mathfrak{A}, \mathfrak{M})$ *to* $\mathfrak{N}$. *Then the functor* F *has a right adjoint if and only if it is right complete and its right localisation has a local right adjoint; in this case the right adjoint functor of* F *is exactly this functor* $lR\,\mathrm{ad}[R\,\mathrm{loc}(F)]$.

Proof. Let us consider F with a right adjoint H. It is clear that the functor F is right complete. By Theorem 6.1 the functor H must have a local left adjoint, and the left adjoint functor F of H is equal to $R\,\mathrm{glo}[lL\,\mathrm{ad}(H)]$. Consequently, the functor $R\,\mathrm{loc}(F)$ is equal to $R\,\mathrm{loc}[R\,\mathrm{glo}[lL\,\mathrm{ad}(H)]]$ or to $lL\,\mathrm{ad}(H)$ by Theorem 4.4, a local left adjoint being always right complete. In other words, the functor $R\,\mathrm{loc}(F)$ has a local right adjoint.

Now let us suppose that F is a right complete functor and that its right localisation has a local right adjoint H. Let us apply Proposition 5.2 to the functor $R\,\mathrm{loc}(F)$. The functors $R\,\mathrm{glo}[R\,\mathrm{loc}(F)]$ and $lR\,\mathrm{ad}[R\,\mathrm{loc}(F)]$ are adjoint. By Theorem 4.3 the first one is equal to F and the second one is equal to H by definition. Thus the functor F has a right adjoint; this functor is equal to $lR\,\mathrm{ad}[R\,\mathrm{loc}(F)]$, and the theorem is completely proved.

Remark 6.3. For a fixed small category $\mathfrak{A}$, the hypothesis "$\mathfrak{M}$ and $\mathfrak{N}$ are right complete categories and F is a right complete functor" is too strong. We only use the existence and preservation of special right roots: right root of type $\hat{\mathfrak{A}}$, and direct sum of type $\mathrm{Hom}_{\mathfrak{A}}(\bar{A}, A)$.

7. Generalization. Let $\mathfrak{M}$ and $\mathfrak{N}$ be two right and left complete categories, and $\mathfrak{A}$ and $\mathfrak{B}$ two small categories. By means of Theorems 6.1 and 6.2 and their duals (*) we can completely solve the problem of knowing whether a functor from $(\mathfrak{A}, \mathfrak{M})$ to $(\mathfrak{B}, \mathfrak{N})$ has a right or left adjoint. Once more we use right localisations, etc.: for instance, the right localisation of a functor F from $(\mathfrak{A}, \mathfrak{M})$ to $(\mathfrak{B}, \mathfrak{N})$ is a functor from $\mathfrak{M}$ to $(\mathfrak{A}^d \times \mathfrak{B}, \mathfrak{N})$; it corresponds to a functor from $\mathfrak{M}$ to $(\mathfrak{A}^d, (\mathfrak{B}, \mathfrak{N}))$, the right localisation of F for the small category $\mathfrak{A}$, the right complete category $\mathfrak{M}$ and the right complete category $(\mathfrak{B}, \mathfrak{N})$.

By means of Theorem 6.2 and 6.1* the following result, generalizing these theorems, can be proved.

THEOREM 7.1. *Let $\mathfrak{M}$ and $\mathfrak{N}$ be two right and left complete categories, $\mathfrak{A}$ and $\mathfrak{B}$ two small categories, and F a functor from $(\mathfrak{A}, \mathfrak{M})$ to $(\mathfrak{B}, \mathfrak{N})$. Then the functor F has a right adjoint if and only if it is right complete and its right localisation has a local right adjoint as a functor from $\mathfrak{M}$ to $(\mathfrak{A}^d \times \mathfrak{B}, \mathfrak{N})$.*

Proof. By means of Theorem 6.2, the functor F has a right adjoint if and only if it is right complete and its right localisation has a local right adjoint (local with respect to $\mathfrak{A}^d$ alone). But by means of Theorem 6.1* a functor from $\mathfrak{M}$ to $(\mathfrak{A}^d \times \mathfrak{B}, \mathfrak{N})$ has a local right adjoint (local with respect to $\mathfrak{A}^d$ alone) if and only if it has a local right adjoint (local with respect to $\mathfrak{A}^d \times \mathfrak{B}$). The theorem is proved.

Remark 7.2. Let us summarize the theory by a self-dual sentence. The right localisation of the left functor and the left localisation of the right functor of a pair of adjoint functors are locally adjoint functors.

III. Examples of Adjoint Functors.

In this part we give several examples of adjoint functors between categories of functors.

8. Functors between small categories. This example is a generalization of the construction "inverse image of a presheaf": see below. The result is the following.

PROPOSITION 8.1. *Let $\mathfrak{M}$ be a right complete category, $\mathfrak{A}$ and $\mathfrak{B}$ two small categories, and F a functor from $\mathfrak{A}$ to $\mathfrak{B}$. Then the functor $(F, \mathfrak{M})$ from $(\mathfrak{B}, \mathfrak{M})$ to $(\mathfrak{A}, \mathfrak{M})$ has a left adjoint.*

Proof. We apply Theorem 6.1 with the categories $(\mathfrak{B}, \mathfrak{M})$ and $\mathfrak{M}$ and the small category $\mathfrak{A}$. We have to prove that for any object A of $\mathfrak{A}$ the functor $\Gamma_A \circ (F, \mathfrak{M})$ has a left adjoint. But this functor is equal to the functor Γ_{FA} from $(\mathfrak{B}, \mathfrak{M})$ to $\mathfrak{M}$, which has a left adjoint according to Lemma 5.1.

Remark 8.2. Proposition 8.1 has an inverse. Let $\mathfrak{M}$ be a category. If for any functor between small categories $(F\colon \mathfrak{A} \to \mathfrak{B})$, the corresponding functor $(F, \mathfrak{M})$ has a left adjoint, then the category $\mathfrak{M}$ is right complete. Think of the case with a small category $\mathfrak{B}$ consisting of a unique morphism.

Example 8.3. Let us consider a continuous map f from Y to X. The dual category of the category of all open sets of X is called $\mathfrak{A}$, and the dual category of the category of all open sets of Y is called $\mathfrak{B}$. Further $F(A)$ is equal to $f^{-1}(A)$. Let $\mathfrak{M}$ be a right complete category. The category $(\mathfrak{A}, \mathfrak{M})$ is the category of all presheaves of X with values in $\mathfrak{M}$, and the category $(\mathfrak{B}, \mathfrak{M})$ is the category of all presheaves of Y with values in $\mathfrak{M}$. The functor $(F, \mathfrak{M})$ is the construction "direct image of a presheaf," and its left adjoint is the construction "inverse image of a presheaf": see [4], pp. 46-49.

Example 8.4. Let $\mathfrak{M}$ be a *right complete Abelian* category and $\mathfrak{B}$ a small category. *If* $\mathfrak{M}$ *has sufficiently many projective objects, then* $(\mathfrak{B}, \mathfrak{M})$ *has sufficiently many projective objects.* The proof is the following. Let us consider the trivial category $\mathfrak{A}$ of all objects and identities of $\mathfrak{B}$ with the injection functor F from $\mathfrak{A}$ to $\mathfrak{B}$. The functor $\Phi = (F, \mathfrak{M})$ has the two following properties. First of all a sequence in $(\mathfrak{B}, \mathfrak{M})$ is exact if and only if its image in $(\mathfrak{A}, \mathfrak{M})$ is exact. Furthermore, by Proposition 8.1 the functor Φ has a left adjoint Ψ. Since the category $(\mathfrak{A}, \mathfrak{M})$ is a product of copies of $\mathfrak{M}$, it has sufficiently many projective objects. Now let us consider an object X of $(\mathfrak{B}, \mathfrak{M})$ and an epimorphism $P \xrightarrow{\pi} \Phi X$ of $(\mathfrak{A}, \mathfrak{M})$, P being projective. We have two interesting morphisms:

$$\begin{tikzcd}\Psi P \arrow[r, "\Psi\pi"] & \Psi\Phi X \arrow[r, "v"] & X\end{tikzcd}$$

the second one corresponding to the identity morphism of ΦX. The first morphism is an epimorphism, for the functor Ψ is right exact, as is every left adjoint. Since Ψ and Φ are adjoint, the morphism Φv is an epimorphism; consequently, so is the morphism v. Finally, the object ΨP is projective, for the functor $\operatorname{Hom}(\Psi P, \cdot)$ is equivalent to the functor $\operatorname{Hom}(P, \Phi\cdot)$, which is exact. The proposition is proved.

9. The category of all sets. The category of all sets and maps has the following property: each object is a direct sum $\sum P_i$ where P_i is equal to a fixed object P (the set P consists of a unique element). This fact has the following consequence.

PROPOSITION 9.1. *Let* $\mathfrak{M}$ *be a right complete category,* $\mathfrak{B}$ *a small category, and* $\mathfrak{E}$ *the category of all sets and maps. Then a functor from* $(\mathfrak{B},\mathfrak{E})$ *to* $\mathfrak{M}$ *has a right adjoint if and only if it is right complete. Furthermore, we may speak of the category of all right complete functors from* $(\mathfrak{B},\mathfrak{E})$ *to* $\mathfrak{M}$ *and of all natural transformations. The category is equivalent to* $(\mathfrak{B},\mathfrak{M})$.

We have to prove that the class of all natural transformations between two right complete functors from $(\mathfrak{B},\mathfrak{E})$ to $\mathfrak{M}$ is a set. For an application of Proposition 9.1, see [6], p. 312.

Proof. We apply Theorem 6.2. Let T be a right complete functor from $(\mathfrak{B},\mathfrak{E})$ to $\mathfrak{M}$. We have to prove that its right localization has a local right adjoint. In other words, for any object B of $\mathfrak{B}$, the functor $\Gamma_B \circ R\operatorname{loc}(T)$ from $\mathfrak{E}$ to $\mathfrak{M}$ must have a right adjoint. This functor is right complete (see 2.1). Since any right complete functor from $\mathfrak{E}$ to $\mathfrak{M}$ has a right adjoint, the first part of the proposition is proved.

The class of all right complete functors from $(\mathfrak{B},\mathfrak{E})$ to $\mathfrak{M}$ is equivalent to the class of all right complete functors from $\mathfrak{E}$ to $(\mathfrak{B}^d,\mathfrak{M})$: see Theorems 4.3 and 4.4. This last class is equivalent to $(\mathfrak{B}^d,\mathfrak{M})$. The second part of the proposition is proved.

10. Zero functors. Let $\mathfrak{P}$ and $\mathfrak{Q}$ be two categories with zero morphisms. A functor from $\mathfrak{P}$ and $\mathfrak{Q}$ is a *zero functor* if and only if the image of a zero morphism of $\mathfrak{P}$ is a zero morphism of $\mathfrak{Q}$.

PROPOSITION 10.1. *Let* $\mathfrak{M}$ *be a right complete category with zero morphisms,* $\mathfrak{A}$ *a small category with zero morphisms, and* $(\mathfrak{A},\mathfrak{M})^0$ *the full subcategory of* $(\mathfrak{A},\mathfrak{M})$ *of all zero functors from* $\mathfrak{A}$ *to* $\mathfrak{M}$. *Then the injection functor of* $(\mathfrak{A},\mathfrak{M})^0$ *into* $(\mathfrak{A},\mathfrak{M})$ *has a left adjoint.*

Proof. We apply Theorem 6.1 with the injection functor H. We have to prove that, for any object $\bar{A}$ of $\mathfrak{A}$, the functor $\Gamma_{\bar{A}} \circ H$ has a left adjoint. A certain functor $T_{\bar{A}}$ from $\mathfrak{M}$ to $(\mathfrak{A},\mathfrak{M})^0$ is very similar to the left adjoint of $\Gamma_{\bar{A}}$ (see Lemma 5.1), and is defined as follows. To an object $M \times A$ of $\mathfrak{M} \times \mathfrak{A}$ there corresponds the direct sum $\sum M_i$ where M_i is equal to M and the index set to the set of all non zero morphisms from $\bar{A}$ to A. To a morphism $\mu \times A$ of $\mathfrak{M} \times \mathfrak{A}$ there corresponds the morphism $\sum \mu_i$ where μ_i is equal to μ and the index set to the set of all non zero morphisms from $\bar{A}$ to A. To a morphism $M \times \alpha$ of $\mathfrak{M} \times \mathfrak{A}$ there corresponds the following morphism. This morphism, composed with the canonical injection of M corresponding to a non zero morphism ω, is equal to the canonical injection of M corresponding to the morphism $\alpha\omega$ if the latter is a non zero morphism, and is a zero morphism if $\alpha\omega$ is a zero morphism. It is easily proved that the functors T_A^- and $\Gamma_A^- \circ H$ are adjoint. Thus the proof of the proposition is settled.

A consequence of this proposition is the following result.

COROLLARY 10.2. *Let* $\mathfrak{M}$ *be a right complete category with zero morphisms,* $\mathfrak{A}$ *a small category with zero morphisms,* $(\mathfrak{A},\mathfrak{M})^0$ *the full subcategory of* $(\mathfrak{A},\mathfrak{M})$ *of all zero functors from* $\mathfrak{A}$ *to* $\mathfrak{M}$ *and* $\mathfrak{P}$ *another category. Then a functor from* $\mathfrak{P}$ *to* $(\mathfrak{A},\mathfrak{M})^0$ *has a left adjoint if and only if the corresponding functor from* $\mathfrak{P}$ *to* $(\mathfrak{A},\mathfrak{M})$ *has a left adjoint.*

This is an obvious consequence of the two following facts: the subcategory is full and the injection functor has a left adjoint.

11. A general procedure. Let $\mathfrak{M}$ be a right and left complete category, $\mathfrak{A}$ and $\mathfrak{B}$ two small categories, and H a functor from the product category $\mathfrak{B}^d \times \mathfrak{A}$ to the category of all sets and maps. Let us consider the definition of the right localisation in Paragraph 2. We replace the functor Hom from $\mathfrak{A}^d \times \mathfrak{A}$ to $\mathfrak{E}$ by the functor H from $\mathfrak{B}^d \times \mathfrak{A}$ to $\mathfrak{E}$. Thus we get a functor ${}^*H$ from $\mathfrak{M}$ to $(\mathfrak{B}^d \times \mathfrak{A}, \mathfrak{M})$. If we proceed dually (direct products instead of direct sums) we get another functor H^* from $\mathfrak{M}$ to $(\mathfrak{B} \times \mathfrak{A}^d, \mathfrak{M})$. These two functors are locally adjoint since the following equality holds:

$$\mathrm{Hom}_{\mathfrak{M}}(\sum_E \bar{M}, M) = \mathrm{Hom}_{\mathfrak{M}}(\bar{M}, \prod_E M).$$

By Theorem 7.1 and Remark 7.2, *the right globalisation of* ${}^*H$ *and the left globalisation of* H^* *are adjoint functors between the categories* $(\mathfrak{A},\mathfrak{M})$ *and* $(\mathfrak{B},\mathfrak{M})$. This construction is of course natural with respect to the right and left complete category $\mathfrak{M}$.

Example 11. 1. A pair of adjoint functors one can construct by this procedure appears in Paragraph 8. The set $H(B,A)$ is equal to the set $\mathrm{Hom}_{\mathfrak{B}}(B,F(A))$. We get the functor $(F,\mathfrak{M})$ and its right adjoint.

12. Dold-Kan adjoint functors. Let $\mathfrak{M}$ be a (complete) Abelian category and $\mathfrak{A}$ and $\mathfrak{B}$ the two following small categories ($\mathfrak{A}$ with zero morphisms). The set of all objects of $\mathfrak{A}$ is equal to the set of all integers $n \geqq 0$. The set $\mathrm{Hom}([m],[n])$ consists of one element θ if m is different from n or $n+1$, of two elements θ and ι if m is equal to n, and of two elements θ and δ if m is equal to $n+1$; the θ's are zero morphisms, the ι's are identity morphisms, and the $\delta\circ\delta$'s are of course zero morphisms. The set of all objects of $\mathfrak{B}$ is equal to the set of all integers $n \geqq 0$. The set $\mathrm{Hom}(\{m\},\{n\})$ consists of all non decreasing maps of the ordered set $(0,1,\cdots,n)$ into the ordered set $(0,1,\cdots,m)$: the composition law of maps yields the composition law of $\mathfrak{B}$.

The category $(\mathfrak{A},\mathfrak{M})^0$ is the category of all chains of $\mathfrak{M}$, and the category $(\mathfrak{B},\mathfrak{M})$ is the category of all semi-simplicial objects of $\mathfrak{M}$. There exists a classical functor N from $(\mathfrak{B},\mathfrak{M})$ to $(\mathfrak{A},\mathfrak{M})^0$. For any functor K from $\mathfrak{B}$ to $\mathfrak{M}$ and any object $[n]$ of $\mathfrak{A}$, the object $NK[n]$ of $\mathfrak{M}$ is equal to the intersection of the kernels of the morphisms $K(d_n{}^i)$ from $K\{n\}$ to $K\{n-1\}$ with $0 < i \leqq n$ (the morphism $K(d_n{}^i)$ is the i-th face operator in dimension n. By restriction the morphism $K(d_n{}^0)$ yields the morphism $NK(\delta)$ in dimension n.

Proposition 12.1. *Let $\mathfrak{M}$ be a (complete) Abelian category. Then the canonical functor N from the category of all semi-simplicial objects of $\mathfrak{M}$ to the category of all chains of $\mathfrak{M}$ has a left adjoint.*

Proof. We apply Corollary 10.2 and Theorem 7.1*. We have to prove that the functor N from $(\mathfrak{B},\mathfrak{M})$ to $(\mathfrak{A},\mathfrak{M})$ is left complete and that its left localisation has a local left adjoint. The intersection of kernels is a left root; consequently, N is a left complete functor. Furthermore, the left localisation of N is a very simple functor. For any objects A of $\mathfrak{A}$ and B of $\mathfrak{B}$ to an object M of $\mathfrak{M}$ there corresponds a finite direct sum $\sum M$, the index set depending on A and B. A functor "finite direct sum" from an Abelian category to the same category is its own adjoint. Consequently, the proposition is proved.

Remark 12.2. In fact we only use finite roots. It can be seen that the completeness property of the Abelian category is not necessary. Furthermore, the categories of all chains of $\mathfrak{M}$ and of all semi-simplicial objects of $\mathfrak{M}$ are equivalent by means of the functor N and its left adjoint: see [3].

13. The geometric realization of a semi-simplicial complex. In Paragraph 9 we studied functors from $(\mathfrak{B},\mathfrak{E})$ to $\mathfrak{M}$. Here we went to study a well-known functor from $\mathfrak{M}$ to $(\mathfrak{B},\mathfrak{E})$—the category $\mathfrak{M}$ being that of all topological spaces, and the small category $\mathfrak{B}$ being that of Paragraph 12. The result is the following proposition, see [7].

Proposition 13.1. *The functor: singular complexes, from the category of all topological spaces to the category of all semi-simplicial complexes has a left adjoint, the functor: geometric realization.*

Proof. We apply once more Theorem 6.1. Let us call Δ_n the n-th geometric standard simplex. We have to prove that the functor by which to a topological space X there corresponds the set of all continuous maps of Δ_n into X has a left adjoint. It is of course the functor by which to a set E there corresponds the topological space $\sum \Delta_n = E \times \Delta_n$, the discrete topology being given to E. Thus the proposition is proved.

BATTELLE MEMORIAL INSTITUTE GENEVA.

REFERENCES.

[1] D. M. Kan, "Adjoint functors," *Transactions of the American Mathematical Society*, vol. 87 (1958), pp. 294-329.

[2] ———, "Functors involving c. s. s. complexes," *Transactions of the American Mathematical Society*, vol. 87 (1958), pp. 330-346.

[3] A. Dodd and D. Puppe, "Homologie nicht-additiver Funktoren," *Annales de l'Institut Fourier*, vol. 11 (1961), pp. 201-312.

[4] J. W. Gray, *Sheaves with values in a category*, Mimeographed notes, Columbia University.

[5] A. Heller and K. A. Rowe, "On the category of sheaves," *American Journal of Mathematics*, vol. 84 (1962), pp. 205-216.

[6] M. André, "Homotopy groups in categories," *Mathematische Zeitschrift*, vol. 87 (1965), pp. 299-313.

[7] J. Milnor, "The geometric realization of a semisimplicial complex," *Annals of Mathematics*, vol. 65 (1957), pp. 377-389.

REPRESENTATIONS INDUCED FROM COMPACT SUBGROUPS.*

By Adam Kleppner.

In this paper we shall explicitly describe a dense subspace of the space of intertwining operators for two representations induced by continuous unitary representations of compact subgroups of a locally compact group, compute the intertwining number of these representations, and from these results deduce necessary and sufficient conditions for irreducibility and disjointness. For the case of representations induced from compact normal subgroups we shall obtain certain other kinds of information. Because at one point it becomes necessary to consider projective representations and because the treatment of projective representations throughout offers essentially no more difficulty than that of ordinary representations, we shall in fact present these results for projective representations. We shall in general employ the terminology and definitions of [7], with two exceptions. What Mackey calls a σ^{-1} representation we call a σ representation, and we shall always suppose the multipliers σ are *normalized*, that is, we suppose $\sigma(x,x)^{-1} = 1$, all x, which implies $\sigma(x,y)^{-1} = \sigma(y^{-1},x^{-1})$, all x, y; see [5] for further details. As is usual, *representation* will always mean *continuous unitary representation.*

Let σ be a multiplier on the locally compact group $\mathfrak{G}$ and let C be the group of complex numbers of modulus 1. With multiplication on $C \times \mathfrak{G}$ defined by: $(\lambda, x)(\gamma, y) = (\lambda\gamma\sigma(x,y), xy)$, $C \times \mathfrak{G}$ becomes a group. If we provide $C \times \mathfrak{G}$ with the product of right Haar measures on C and $\mathfrak{G}$ and apply Weil's theorem (the converse to the existence of Haar measure) we provide $C \times \mathfrak{G}$ with a topology compatible with the group structure for which C is a closed normal subgroup with a quotient topologically isomorphic to G. Thus $C \times \mathfrak{G}$ becomes an extension of a compact group by a locally compact group and by a theorem of Gleason it is itself a locally compact group. This group is denoted by $\mathfrak{G}^{\sigma}$ and was introduced by Mackey in the separable case. In $L^2(\mathfrak{G}^{\sigma})$, formed with respect to right Haar measure, let

---

* This work was partially supported by the National Science Foundation under Grant NSF-G25229.

Received July 28, 1964.
Revised March 8, 1966.

$\boldsymbol{H}_0$ be the subspace of all f which satisfy $f(\lambda,x)=\lambda f(1,x)$, $\lambda\in C$. The map $f\to Ff$ where

$$Ff(\lambda,x)=\lambda\int_C \gamma^{-1}f(\gamma,x)\,d\gamma$$

is the projection of $L^2(\mathcal{G}^\sigma)$ on $\boldsymbol{H}_0$ ($d\gamma$ is normalized Haar measure on C). If for $f\in L^2(\mathcal{G})$ (also formed with right Haar measure) we set $f'(\lambda,x)=\lambda f(x)$, then $f\to f'$ is an isometry of $L^2(\mathcal{G})$ with $\boldsymbol{H}_0$. Let R be the right regular representation of $\mathcal{G}^\sigma$ in $L^2(\mathcal{G}^\sigma)$. The right regular σ representation R^σ of $\mathcal{G}$ in $L^2(\mathcal{G})$ is equivalent to the restriction to $\mathcal{G}$ (identified to the set of all $(1,x)\in\mathcal{G}^\sigma$) of the subrepresentation of R on $\boldsymbol{H}_0$. In fact,

$$(R_sf)'(\lambda,x)=\lambda R^\sigma{}_sf(x)=\lambda\sigma(x,s)f(xs)=R_{1,s}f'(\lambda,x).$$

Let $\mathcal{A}(R)$ be the space of all Haar measurable essentially bounded complex valued functions with compact support on $\mathcal{G}^\sigma$. To each $\phi\in\mathcal{A}(R)$ corresponds a convolution operator V_ϕ on $L^2(\mathcal{G}^\sigma)$ defined by:

$$V_\phi f(\lambda,x)=\int_G\int_C\phi(\gamma,y)f(\gamma\lambda\sigma(y,x),yx)\,d\gamma dy$$

(dy is right Haar measure on $\mathcal{G}$). The operators V_ϕ are dense in $\boldsymbol{R}(R)$ for the weak operator topology (we write $\boldsymbol{R}(R)$ for $\boldsymbol{R}(R,R)$ of [7]). Thus the operators $FV_\phi F$ transferred to $L^2(\mathcal{G})$ by the inverse of the map $f\to f'$ are in $\boldsymbol{R}(R^\sigma)$ and are in fact weakly dense in $\boldsymbol{R}(R^\sigma)$. Carrying out the computations involved we find V_ϕ corresponds to the operator $S_{\phi'}$, where

$$\phi'(y)=\int_C\gamma\phi(\gamma,y)\,d\gamma$$

and

$$S_{\phi'}f(x)=\int_G\phi'(y)\sigma(y,x)f(yx)\,dy.$$

The function ϕ' is a measurable essentially bounded function with compact support on $\mathcal{G}$; let $\mathcal{A}(R^\sigma)$ be the space of all such functions. The map $\phi\to\phi'$ is a linear map of $\mathcal{A}(R)$ onto $\mathcal{A}(R^\sigma)$ and we conclude that the space of all convolution operators S_ϕ, $\phi\in\mathcal{A}(R^\sigma)$, is weakly dense in $\boldsymbol{R}(R^\sigma)$.

Let $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ be finite dimensional complex Hilbert spaces and let R^i be the right regular σ representation of $\mathcal{G}$ in $L^2(\mathcal{G},\boldsymbol{H}_i)$. Provide $\mathcal{L}(\boldsymbol{H}_1,\boldsymbol{H}_2)$ with the Hilbert-Schmidt norm and let $\mathcal{A}(R^1,R^2)$ be the space of all measurable essentially bounded functions with compact support $\alpha\colon\mathcal{G}\to\mathcal{L}(\boldsymbol{H}_1,\boldsymbol{H}_2)$. For each $\alpha\in\mathcal{A}(R^1,R^2)$ and $f\in L^2(\mathcal{G},\boldsymbol{H}_1)$ set

$$T_\alpha f(x) = \int_G \alpha(y) f(yx) \sigma(y,x)\, dy.$$

Then $T_\alpha \in \mathbf{R}(R^1, R^2)$ and further the space of all operators T_α, $\alpha \in \mathcal{A}(R^1, R^2)$, is weakly dense in $\mathbf{R}(R^1, R^2)$. To confirm this last statement let $T \in \mathbf{R}(R^1, R^2)$, let $u_1, \cdots, u_m$ and $v_1, \cdots, v_n$ be orthonormal bases in $\mathbf{H}_1$ and $\mathbf{H}_2$ respectively. For each i, j and $f, g \in L^2(\mathfrak{G})$ set $(T_{ij} f, g) = (T(f \otimes u_i), g \otimes v_j)$, where $f \otimes u_i$ is the function $x \rightarrow f(x) u_i$ and similarly for $g \otimes v_j$. $T_{ij} \in \mathbf{R}(R^\sigma)$ and by the preceding results T_{ij} is a weak limit of convolution operators T_{ijk} on $L^2(\mathfrak{G})$, where $T_{ijk} f(x) = \int \phi_{ijk}(y) f(yx) \sigma(y,x)\, dy$ and $\phi_{ijk} \in \mathcal{A}(R^\sigma)$. We can write each $f \in L^2(\mathfrak{G}, \mathbf{H}_1)$ as $f = \sum f_i \otimes u_i$, where $f_i(x) = (f(x), u_i)$, and similarly for $g \in L^2(\mathfrak{G}, \mathbf{H}_2)$. Thus

$$(Tf, g) = \sum (T_{ij} f_i, g_j) = \lim_k \sum (T_{ijk} f_i, g_j).$$

For each $x \in \mathfrak{G}$ define $\alpha_k(x) \in \mathfrak{L}(\mathbf{H}_1, \mathbf{H}_2)$ by: $(a_k(x) u_i, v_j) = \phi_{ijk}(x)$. Then $\alpha_k \in \mathcal{A}(R^1, R^2)$, $(T_{\alpha_k} f, g) = \sum (T_{ijk} f_i, g_j)$ and $\mathrm{T} = \lim_k T_{\alpha_k}$.

Let L be a σ representation of the compact subgroup G_1 of $\mathfrak{G}$ on the finite dimensional Hilbert space $\mathbf{H}(L)$ and let R^L be the right regular σ representation of $\mathfrak{G}$ in $L^2(\mathfrak{G}, \mathbf{H}(L))$. The induced σ representation U^L is equivalent to the subrepresentation of R^L on the subspace $\mathbf{H}(U^L)$ of $L^2(\mathfrak{G}, \mathbf{H}(L))$ consisting of all f which satisfy $f(\xi x) = \sigma(\xi, x)^{-1} L_\xi f(x)$, all $\xi \in G_1$, the equivalence arising from the fact that, assuming as always the Haar measure on G_1 is normalized, the integral over $\mathfrak{G}$ of a function constant on the right G_1 cosets is the same as its integral over $\mathfrak{G}/G_1$ with respect to the invariant measure determined by Haar measures on $\mathfrak{G}$ and G_1. Because σ is normalized we can define E_L: $L^2(\mathfrak{G}, \mathbf{H}(L)) \rightarrow \mathbf{H}(U^L)$ by:

$$E_L f(x) = \int_{G_1} \sigma(\xi, x) L_\xi^{-1} f(\xi x)\, d\xi$$

E_L is the projection of $L^2(\mathfrak{G}, \mathbf{H}(L))$ onto $\mathbf{H}(U^L)$. If M is a σ representation of the compact subgroup G_2 of $\mathfrak{G}$ we define $R^M, \mathbf{H}(U^M)$ and E_M similarly.

THEOREM 1. *Let σ be a normalized multiplier on the locally compact group $\mathfrak{G}$ and let L and M be finite dimensional representations of the compact subgroups G_1 and G_2 respectively. Let $\mathcal{A}(U^L, U^M)$ be the set of all measurable essentially bounded function with compact support*

$$\alpha\colon \mathfrak{G} \rightarrow \mathfrak{L}(\mathbf{H}(L), \mathbf{H}(M))$$

which satisfy:

$$\alpha(\xi y\eta)=\sigma(\xi,y)\sigma(y,\eta)M_{\eta}^{-1}\alpha(y)L_{\xi}^{-1},\quad \xi\in G_1,\quad \eta\in G_2.$$

For all $\alpha\in\mathcal{A}(U^L,U^M)$, *the convolution operator* T_α *is in* $\mathbf{R}(U^L,U^M)$ *and the space of operators* T_α, $\alpha\in\mathcal{A}(U^L,U^M)$, *is weakly dense in* $\mathbf{R}(U^L,U^M)$.

Proof. For each $\alpha\in\mathcal{A}(R^L,R^M)$, $E_MT_\alpha E_L\in\mathbf{R}(U^L,U^M)$ and the space of these operators is weakly dense in $\mathbf{R}(U^L,U^M)$. For each $f\in L^2(\mathcal{G},\mathbf{H}(L))$ we have

$$\begin{aligned}E_MT_\alpha E_Lf(x)\\&=\int_{G_2}\int_{\mathcal{G}}\int_{G_1}\sigma(\eta,x)\sigma(\xi,y\eta x)\sigma(y,\eta x)M_\eta^{-1}\alpha(y)L_\xi^{-1}f(\xi y\eta x)\,d\xi\,dy\,d\eta\\&=\int_{\mathcal{G}}\int_{G_2}\int_{G_1}\sigma(\xi,y)^{-1}\sigma(\xi y,\eta)^{-1}M_\eta\alpha(\xi y\eta)L_\xi\,d\xi\,d\eta\,f(yx)\,dy\\&=T_{\alpha^0}f(x)\end{aligned}$$

where

$$\alpha^0(y)=\int_{G_2}\int_{G_1}\sigma(\xi,y)^{-1}\sigma(\xi y,\eta)^{-1}M_\eta\alpha(\xi y\eta)L_\xi\,d\xi\,d\eta.$$

The map $\alpha\to\alpha^0$ is the projection of $\mathcal{A}(R^L,R^M)$ onto $\mathcal{A}(U^L,U^M)$ and this proves the theorem.

In the actual computations involved in the proof strong use is made of the fact that σ is normalized and that the restrictions to G_1 and G_2 of the modular function of $\mathcal{G}$ are trivial.

The conclusion of this theorem remains true when $\mathcal{A}(U^L,U^M)$ is replaced by its subspace of continuous functions, or in the case of a Lie group, by its subspace of C^∞ functions, since operators of convolutions by continuous or C^∞ functions with compact support are also weakly dense in $\mathbf{R}(R^L,R^M)$.

We now apply the techniques of § 4 of [4] to compute the dimension of $\mathcal{A}(U^L,U^M)$. Since the map $\alpha\to T_\alpha$ is an injective linear map of $\mathcal{A}(U^L,U^M)$ (the elements of $\mathcal{A}(U^L,U^M)$ being regarded as classes of functions equal almost everywhere) onto a dense subspace of $\mathbf{R}(U^L,U^M)$, $\mathcal{A}(U^L,U^M)$ and $\mathbf{R}(U^L,U^M)$ have the same dimension, when we identify all infinite cardinals to ∞.

For each $x\in\mathcal{G}$ the map $\xi\to\sigma(x,\xi)^{-1}\sigma(x\xi,x^{-1})^{-1}L_{x\xi x^{-1}}$ is a σ representation of $x^{-1}G_1x$ (see [7], corollary to Lemma 4.2) which we call L^x. If $\xi\in G_1$ then $L^{\xi x}$ is equivalent to L^x; in fact, for all $\xi'\in G_1$, $L_\xi L^x{}_{\xi'}=L^{x\xi}{}_{\xi'}L_\xi$. Define the representations M^y similarly and let $\mathbf{R}(L,M,x,y)$ denote the space of intertwining operators for the restrictions to $G_{x,y}=x^{-1}G_1x\cap y^{-1}G_2y$ of the representations L^x and M^y. As is easily verified, $\mathbf{R}(L,M,x,y)=\mathbf{R}(L,M,xy^{-1},e)$ and $M_\eta\mathbf{R}(L,M,x,y)L_\xi^{-1}=\mathbf{R}(L,M,\xi x,\eta y)$. Hence $\dim\mathbf{R}(L,M,x,y)$ depends only on the $G_1 : G_2$ double coset a to which xy^{-1} belongs and may be denoted by $\boldsymbol{I}(L, M, a)$. We shall show that $(\boldsymbol{R}(L, M, x, y))_{x,y \in \mathfrak{G} \times \mathfrak{G}}$ is a *closed* family of subspaces of $\mathcal{L}(\boldsymbol{H}(L), \boldsymbol{H}(M))$ in the sense of § 4 of [4].

Let π be the canonical map of $\mathfrak{G}^\sigma$ on $\mathfrak{G} = \mathfrak{G}^\sigma / C$. $\pi^{-1}(G_1)$ is a compact subgroup of $\mathfrak{G}^\sigma$ and is the subgroup of all (λ, x), $x \in G_1$. The topology induced on $\pi^{-1}(G_1)$ by $\mathfrak{G}^\sigma$ is the same as that induced by $(G_1)^\sigma$ and we can identify $(G_1)^\sigma$ with a compact subgroup of $\mathfrak{G}^\sigma$. More generally, we note that

$$\pi^{-1}(x^{-1} G_1 x) = (\lambda, x)(G_1)^\sigma (\lambda, x)^{-1} = (x^{-1} G_1 x)^\sigma, \text{ all } \lambda \in C,$$

and

$$\pi^{-1}(G_{x,y}) = (G_{x,y})^\sigma = (\lambda, x)(G_1)^\sigma (\lambda, x)^{-1} \cap (\gamma, y)(G_2)^\sigma (\gamma, y)^{-1}.$$

This group we denote by $G_{x,y}{}^\sigma$. Let L^0 be the (ordinary) representation $(\lambda, \xi) \to \lambda L_\xi$ of $(G_1)^\sigma$. Then $(L^0)^{\lambda, x} = (L^x)^0$, all λ. We define similarly M^0. Let $\boldsymbol{R}(L^0, M^0, (\lambda, x), (\gamma, y))$ be the space of intertwining operators for the restrictions to $G_{x,y}{}^\sigma$ of $(L^0)^{\lambda, x}$ and $(M^0)^{\gamma, y}$. Obvious modifications of Lemma 9 of [4] show that

$$(\boldsymbol{R}(L^0, M^0, (\lambda, x), (\gamma, y)))_{(\lambda, x), (\gamma, Y) \in \mathfrak{G}^\sigma \times \mathfrak{G}^\sigma}$$

is a closed family of subspaces of $\mathcal{L}(\boldsymbol{H}(L), \boldsymbol{H}(\mathrm{M}))$, that is, the set of all $(\lambda, x), (\gamma, y), T$ such that $T \in \boldsymbol{R}(L^0, M^0, (\lambda, x), (\gamma, y))$ is a closed subset of $\mathfrak{G}^\sigma \times \mathfrak{G}^\sigma \times \mathcal{L}(\boldsymbol{H}(L), \boldsymbol{H}(M))$. The map $(\lambda, x), (\gamma, y), T \to x, y, T$ is then a closed map of $\mathfrak{G}^\sigma \times \mathfrak{G}^\sigma \times \mathcal{L}(\boldsymbol{H}(L), \boldsymbol{H}(M))$ onto $\mathfrak{G} \times \mathfrak{G} \times \mathcal{L}(\boldsymbol{H}(L), \boldsymbol{H}(M))$ and under this map the famiy $(\boldsymbol{R}(L^0, M^0, (\lambda, x), (\gamma, y)))_{(\lambda, x), (\gamma, y) \in \mathfrak{G}^\sigma \times \mathfrak{G}^\sigma}$ is carried onto the family $(\boldsymbol{R}(L^0, M^0, (1, x), (1, y)))_{x,y \in \mathfrak{G} \times \mathfrak{G}}$. But $T \in \boldsymbol{R}(L^0, M^0, (1, x), 1, y))$ if and only if

$$T(L^0)^{(1,x)}_{(\lambda,\eta)} = \lambda T L^x{}_\eta = (M^0)^{(1,y)}_{(\lambda,\eta)} T = \lambda M^y{}_\eta T, \text{ all } (\lambda, \eta) \in G_{x,y},$$

that is, if and only if $T \in \boldsymbol{R}(L, M, x, y)$. Hence, $(\boldsymbol{R}(L, M, x, y))_{x,y \in \mathfrak{G} \times \mathfrak{G}}$ is a closed family of subspaces of $\mathcal{L}(\boldsymbol{H}(L), \boldsymbol{H}(M))$.

Let $\alpha \in \mathfrak{A}(U^L, U^M)$ and set $\alpha'(x, y) = \alpha(xy^{-1})$. Then α' is a measurable essentially bounded function on $\mathfrak{G} \times \mathfrak{G}$ with compact support modulo the diagonal subgroup of $\mathfrak{G} \times \mathfrak{G}$. That α satisfies

$$\alpha(\xi y \eta) = \sigma(\xi, y) \sigma(y, \eta) M_\eta{}^{-1} \alpha(y) L_\xi{}^{-1}, \quad \xi \in G_1, \quad \eta \in G_2,$$

almost all $y \in \mathfrak{G}$, means that for almost all $x, y, \alpha'(x, y) \in \boldsymbol{R}(L, M, x, y)$. Conversely, let α' be a measurable essentially bounded function $\mathfrak{G} \times \mathfrak{G} \to \mathcal{L}(\boldsymbol{H}(L), \boldsymbol{H}(M))$ with compact support modulo the diagonal subgroup of $\mathfrak{G} \times \mathfrak{G}$ such that $\alpha'(x, y) \in \boldsymbol{R}(L, M, x, y)$, almost all x, y. Then the function $\alpha(xy^{-1}) = \alpha'(x, y) \in \mathfrak{A}(U^L, U^M)$. Thus with the notation of [4] $\alpha \to \alpha'$ is an injective map of $\mathfrak{A}(U^L, U^M)$ onto the subspace of $L^\infty(\mu \times \mu, (\boldsymbol{R}(L, M, x, y)))$

consisting of functions with compact support modulo the diagonal subgroup of $\mathfrak{G} \times \mathfrak{G}$ (μ is right Haar measure on $\mathfrak{G}$). Because G_1 and G_2 are compact the discussion of §4 of [4] shows that the dimension of this space is the same as that of its subspace of functions with compact support, and thus the same as the dimension of the subspace of functions with compact support modulo the diagonal subgroup of $\mathfrak{G} \times \mathfrak{G}$. Noting that G_1 and G_2 are regularly related, at least if $\mathfrak{G}$ is separable, we can immediately deduce from §4 of [4] the following intertwining number theorem. We put $\boldsymbol{I}(U^L, U^M) = \dim \boldsymbol{R}(U^L, U^M)$.

THEOREM 2. *Let σ be a normalized multiplier on the separable locally compact group $\mathfrak{G}$ and let L and M be finite dimensional σ representations of the compact subgroups G_1 and G_2 respectively. Let ν be the image in the space of G_1: G_2 double cosets of Haar measure and write $\nu = \nu_1 + \nu_2$, where ν_1 is a purely atomic Borel measure normalized so that the measure of a set is the number of atoms in it and ν_2 is a non-atomic measure. If $\boldsymbol{I}(L, M, a) > 0$ on a set of positive ν_2 measure then $\boldsymbol{I}(U^L, U^M) = \infty$. If $\boldsymbol{I}(L, M, a) = 0$, ν_2 almost everywhere, then*

$$\boldsymbol{I}(U^L, U^M) = \int \boldsymbol{I}(L, M, a)\, d\nu_1(a).$$

COROLLARY 1. *For U^L to be irreducible it is necessary and sufficient that $\boldsymbol{I}(L, L, a) = 0$, ν_2 almost everywhere, and that $\boldsymbol{I}(L, L, a) = 1$ for exactly one atom a.*

COROLLARY 2. *For U^L and U^M to be disjoint it is necessary and sufficient that the representations $L^x \mid G_{x,y}$ and $M^y \mid G_{x,y}$ be disjoint for almost all x, y.*

COROLLARY 3. *Suppose G_1 and G_2 are normal and G_1G_2, which is a compact normal subgroup, is not open. Then $\boldsymbol{I}(U^L, U^M) = 0$ or ∞, and $\boldsymbol{I}(U^L, U^L) = \infty$.*

In fact, in this case the space of G_1: G_2 double cosets is identified with the group $\mathfrak{G}/G_1G_2$ and ν is equivalent to Haar measure on this group. Since G_1G_2 is not open there are no atoms, $\nu_1 = 0$ and the first assertion follows from the theorem. For the second we note that G_1 is not open and $\boldsymbol{I}(U^L, U^L) \geqq 1$ which implies $\boldsymbol{I}(U^L, U^L) = \infty$.

Let G be a compact normal subgroup of $\mathfrak{G}$ and let L be an irreducible σ representation of G. The set of all $x \in \mathfrak{G}$ such that $\boldsymbol{I}(L^x, L) \neq 0$ is a subgroup G_L of $\mathfrak{G}$ containing G called the stability group of L.

COROLLARY 4. *Let L be an irreducible σ representation of the compact normal subgroup G of $\mathfrak{G}$. Then G_L is open.*

If G is open, clearly G_L is open. Suppose G is not open. By Corollary 3, $\boldsymbol{I}(U^L, U^L) = \infty$ and by Theorem 2, $\boldsymbol{I}(L, L, a) > 0$ on a measurable set of positive ν measure. The inverse image in $\mathfrak{G}$ of this set is a measurable set S of positive Haar measure contained in G_L. But then $G_L \supset SS^{-1}$ and SS^{-1} contains a neighborhood of $e \in \mathfrak{G}$.

COROLLARY 5. *With the hypotheses of Corollary 4, suppose further that* $\mathfrak{G}$ *is connected. Then* $G_L = \mathfrak{G}$.

For not necessarily separable Lie groups the disjointness criterion of Corollary 2 was found by Blattner [1], and in [2] he found a somewhat more general, though less exact result, than Theorem 2. If in Theorem 2 the subgroups G_1 and G_2 are open then $G_{x,y}$ is of finite index in both the containing groups and the expression found here for $\boldsymbol{I}(U^L, U^M)$ is the same as that found in [6] and [8]. A direct proof of Corollary 4 can be obtained by observing that $\mathfrak{G}$ acts on the dual of G, which is a discrete Borel space and that the orbits are coset spaces.

Suppose now that $\mathfrak{G}$ is separable, σ is a normalized multiplier on $\mathfrak{G}$ and L is an irreducible σ representation of the compact normal subgroup G of $\mathfrak{G}$. Let G_L be the stability subgroup of L. According to [7], Theorem 8.2, there exists a multiplier τ on G_L and a τ representation M of G_L such that $\tau = \sigma\omega^{-1}$, where ω is a normalized multiplier constant on the G cosets in $\mathfrak{G}$, $M_{\xi x} = L_\xi M_x$, all $\xi \in G$, and $M_x \in \boldsymbol{R}(L^x, L)$, all $x \in G_L$. We shall regard ω as being a multiplier defined on G_L/G and say ω is the *multiplier associated to L*. Let $\alpha \in \mathfrak{A}(U^L)$ (we write $\mathfrak{A}(U^L)$ for the space $\mathfrak{A}(U^L, U^L)$ defined in Theorem 1) and $\xi \in G$. Then $\alpha(\xi^{-1}x) = \alpha(x)L_\xi = \alpha(xx^{-1}\xi^{-1}x) = L^{x^{-1}}{}_\xi\alpha(x)$ and $\alpha(x) \in \boldsymbol{R}(L, L^{x^{-1}})$. In particular, $\alpha(x) = 0$ if $x \notin G_L$. Because L is irreducible, $\boldsymbol{R}(L, L^{x^{-1}})$ consists of scalar multiples of the unitary operator $M_x{}^{-1}$ and $\alpha(x) = \phi_\alpha(x) M_x{}^{-1}$ for some complex number $\phi_\alpha(x)$, $x \in G_L$. ϕ_α is a measurable essentially bounded function with compact support on G_L and from the equations $\phi_\alpha(\xi x) M_{\xi x}{}^{-1} = \alpha(\xi x) = \alpha(x) L_\xi{}^{-1} = \phi_\alpha(x) M_x{}^{-1} L_\xi{}^{-1} = \phi_\alpha(x) M_{\xi x}{}^{-1}$ we conclude that ϕ_α is constant on the G cosets in G_L. Thus ϕ_α defines a measurable essentially bounded function on G_L/G with which we identify it. The function ϕ_α clearly determines α and the convolution operator T_α. Making use of this correspondence we shall prove the following theorem describing $\boldsymbol{R}(U^L)$.

THEOREM 3. *Let* σ *be a normalized multiplier on the separable locally compact group* $\mathfrak{G}$, *let* L *be an irreducible* σ *representation of the compact normal subgroup* G *of* $\mathfrak{G}$, *let* G_L *be the stability group of* L, *let* ω *be the multiplier on* G_L/G *associated to* L *and let* R^ω *be the right regular* ω *representation of* G_L/G. *Then the von Neumann algebras* $\boldsymbol{R}(U^L)$ *and* $\boldsymbol{R}(R^\omega)$ *are isomorphic.*

Proof. We shall show first that $\mathbf{R}(U^L)$ is isomorphic to $\mathbf{R}(N)$ where N is the σ representation of G_L induced by L. Let $\{x_i\}_{i \in I}$ be a set of right coset representatives of G_L in $\mathfrak{G}$ with $x_0 = e$. Let $\mathbf{H}_i$ be the subspace of $\mathbf{H}(U^L)$ consisting of all functions which vanish off $G_L x_i$. Because G_L is open $\mathbf{H}(U^L)$ is the Hilbert space direct sum of the subspaces $\mathbf{H}_i$, and if E_i is the projection on $\mathbf{H}_i$, $E_i T_\alpha = T_\alpha E_i$ for all $\alpha \in \mathcal{A}(U^L)$. Because the T_α are dense in $\mathbf{R}(U^L), E_i \in \mathbf{R}(U^L)'$. The operator $U^L{}_{x_i} U^L{}_{x_j} E_i$ is a partial isometry in $\mathbf{R}(U^L)'$ with initial projection E_i and final projection E_j. Because the $U_x{}^L$, $x \in \mathfrak{G}$, generate $\mathbf{R}(U^L)'$ the set of operators $E_0 U^L{}_x E_0$, $x \in \mathfrak{G}$, generates the reduced algebra $\mathbf{R}(U^L)'_{E_0}$ (terminology of [3]). But $E_0 U^L{}_x E_0 = 0$ if $x \notin G_L$ and by the definition of N, $\mathbf{H}_0 = \mathbf{H}(N)$. Thus $\mathbf{R}(U^L)'_{E_0} = \mathbf{R}(N)'$ and it now follows from [3], Prop. 5, Ch. 1, § 2, that $\mathbf{R}(U^L)'$ is isomorphic to $\mathbf{R}(N)' \otimes \mathcal{L}(L^2(\mathfrak{G}/G_L))$, and $\mathbf{R}(U^L)$ is isomorphic to $\mathbf{R}(N)$. Thus we may now suppose that $\mathfrak{G} = G_L$.

Let I be the identity operator on $\mathbf{H}(L)$. For each $\phi \in \mathcal{A}(R^\omega)$, the operator $S_\phi \otimes I$ is an operator on $L^2(\mathfrak{G}/G) \otimes \mathbf{H}(L) = L^2(\mathfrak{G}/G, \mathbf{H}(L))$ and the von Neumann algebra generated by the operators $S_\phi \otimes I$ is isomorphic to $\mathbf{R}(R^\omega)$. Explicitly, $S_\phi \otimes I$ is the operator

$$((S_\phi \otimes I)f, g) = \int_{\mathfrak{G}/G} \int_{\mathfrak{G}/G} \phi(y)\,(f(yx), g(x))\,\omega(y, x)\,dy\,dx$$

$$= \int_{\mathfrak{G}/G} \int_{\mathfrak{G}/G} \phi(yx^{-1})\,(f(y), g(x))\,\omega(x, y^{-1})\,dy\,dx,$$

$f, g \in L^2(\mathfrak{G}/G, \mathbf{H}(L))$. For each $f \in \mathbf{H}(U^L)$ the function $x \to M_x{}^{-1} f(x)$ is constant on the G cosets and defines a function f' in $L^2(\mathfrak{G}/G, \mathbf{H}(L))$: The map $f \to f'$ is an isometry of $\mathbf{H}(U^L)$ on $L^2(\mathfrak{G}/G, \mathbf{H}(L))$. For each $\alpha \in \mathcal{A}(U^L)$ and $f, g \in \mathbf{H}(U^L)$ we have

$$\begin{aligned}(T_\alpha f, g) &= \int\int \alpha(y)\,(f(yx), g(x))\,\sigma(y, x)\,dy\,dx \\ &= \int\int \alpha(yx^{-1})\,(f(y), g(x))\,\sigma(x, y^{-1})\,dy\,dx \\ &= \int\int \phi_\alpha(yx^{-1})\,(M_y{}^{-1} f(y), M_x{}^{-1} g(x))\,\omega(x, y^{-1})\,dy\,dx.\end{aligned}$$

This last integral is constant on the G cosets and its integral over $\mathfrak{G}/G \times \mathfrak{G}/G$ is the same as its integral over $\mathfrak{G} \times \mathfrak{G}$, from which we conclude

$$(T_\alpha f, g) = ((S_{\phi_\alpha} \otimes I) f', g').$$

The map $T_\alpha \to S_{\phi_\alpha} \otimes I$ is an isomorphism of a dense *subalgebra of $\mathbf{R}(U^L)$ onto a dense *subalgebra of a von Neumann algebra isomorphic to $\mathbf{R}(R^\omega)$. The equation above says this isomorphism is bicontinuous for the weak topologies on these algebras and may be extended to give a bicontinuous isomorphism of $\boldsymbol{R}(U^L)$ onto $\boldsymbol{R}(R^\omega)$.

The analogous theorem for discrete groups was proved in [5], but there the multiplier ω is called ω^{-1}.

University of Maryland,
College Park, Maryland.

REFERENCES.

[1] R. Blattner, "On induced representations," *American Journal of Mathematics*, vol. 83 (1961), pp. 79-98.

[2] ———, "A theorem on induced representations," *Proceedings of the American Mathematical Society*, vol. 13 (1962), pp. 881-884.

[3] J. Dixmier, *Les algèbres d'operateurs dans l'espace Hilbertien*, Gauthier-Villars, Paris, 1957.

[4] A. Kleppner, "Intertwining forms for summable induced representations," *Transactions of the American Mathematical Society*, vol. 112 (1964), pp. 164-183.

[5] ———, "The structure of some induced representations," *Duke Mathematical Journal*, vol. 29 (1962), 555-572.

[6] ———, " On the intertwining number theorem, *Proceedings of the American Mathematical Society*, vol. 12 (1961), pp. 731-733.

[7] G. W. Mackey, "Unitary representations of group extensions I," *Acta Mathematica*, vol. 99 (1958), pp. 265-311.

[8] ———, "On induced representations of groups," *American Journal of Mathematics*, vol. 83 (1951), pp. 576-592.

ORTHOGONAL GROUPS OVER THE INTEGERS OF A LOCAL FIELD I.

By Carl R. Riehm.*

Let $\mathfrak{o}$ be the ring of integers in a non-dyadic local field (the residue class field is therefore finite and of characteristic > 2). Let L be a finitely-generated free $\mathfrak{o}$-module endowed with a symmetric non-degenerate bilinear form—here non-degenerate means non-zero discriminant. The orthogonal group $O(L)$ is the group of isometries of L.

The structure of $O(L)$ has been studied by Klingenberg [3] when L is unimodular (discriminant a unit), and when the ring $\mathfrak{o}$ is, more generally, a local ring. His results are very similar to those of Dieudonné ([1] or [2]) in the case of a field. The present paper gives intial[1] results on the structure of $O(L)$ in the local integral case without the restriction of unimodularity on L, and is independent of [3].

Let $\Omega(L)$ be the commutator subgroup of $O(L)$. The main concern of this paper is to describe the groups contained between $O(L)$ and $\Omega(L)$ and the corresponding factor groups (Theorem 3.3), to give generators for these groups (Theorem 3.4) and to determine their centers (Theorem 3.5), and finally to give alternate characterizations of $\Omega(L)$ (Theorems 3.2 and 3.6).

The results depend on a technique of "reduction mod $\mathfrak{p}$" (§ 2) which makes it possible to "lift" properties of the orthogonal group over finite fields. Partly for this reason, the theory is radically different from the field and local ring theory. For example the "spinor group" $O'(L)$ (as defined in § 2) is always equal to the commutator subgroup.

§ 1. Lattices and Orthogonal Groups.

The ring $\mathfrak{o}$ will be a complete discrete valuation ring with a finite residue class field k of characteristic > 2. The maximal ideal of $\mathfrak{o}$ is $\mathfrak{p}$. For such

---

Received December 3, 1964.

* The research for this paper was supported in part by the National Science Foundation under grants GP-1656 and GP-3986.

[1] In a sequel to this paper I shall show that the normal subgroups of $\Omega(L)$ (with restrictions on L analogous to those in the case of a field) are congruence groups in the sense of my paper "The structure of the symplectic group over a valuation ring" on pp. 106-128 of this volume.

a ring $\mathfrak{o}$ we have the *local square theorem* (63:1, [4]): *if* α *and* β *are elements of* $\mathfrak{o}$ *such that* $\alpha \equiv \beta \bmod \alpha\mathfrak{p}$, *then* $\alpha = \epsilon^2\beta$ *for some unit* ϵ.

Now let L be a *lattice*, i. e. a finitely-generated free $\mathfrak{o}$-module endowed with a non-degenerate symmetric bilinear form, denoted by $(X, Y) \rightarrow X \cdot Y$ where $X, Y \in L$. Thus $X \cdot Y = Y \cdot X$ for all X and Y in L, and $X \cdot L = 0$ implies that $X = 0$. We often write $X \cdot X = X^2$. The dimension (rank) of L is denoted by $\dim L$.

According to § 92, [4], L has an orthogonal basis, i. e. a basis $A_1, \cdots, A_n$ such that $A_i \cdot A_j = 0$ if $i \neq j$. A vector in L is called *pure* if it is contained in some orthogonal basis of L. If L is the direct sum of two submodules M and N which are mutually orthogonal ($M \cdot N = 0$), then we write $L = M \perp N$; N is uniquely determined by M as the submodule M° of L orthogonal to M, and both are lattices. A lattice L is said to be *modular* if it has an orthogonal basis $A_1, \cdots, A_n$ such that $A_1 \cdot L = A_2 \cdot L = \cdots = A_n \cdot L$. Any lattice L has a *Jordan splitting*; that is,

$$L = L_1 \perp L_2 \perp \cdots \perp L_t \tag{1}$$

where each L_i is modular and $L_1 \cdot L_1 \supset L_2 \cdot L_2 \supset \cdots \supset L_t \cdot L_t$ (proper containment). In 91:9 of [4], it is shown that t, the ideals $\mathfrak{s}_i = L_i \cdot L_i$, and $\dim L_i$, $(i = 1, \cdots, t)$, are invariants of L. A pure vector $A \in L$ is said to be *pure of type* i if $A \cdot L = \mathfrak{s}_i$.

Another important fact about modular lattices is the *modular component theorem* (82:15, [4]): *a modular lattice* $M \subseteq L$ *is an orthogonal component of* L (*i. e.* $L = M \perp M^\circ$) *if and only if* $M \cdot M = M \cdot L$.

Two lattices L and K are *isometric* if there is an isomorphism ("isometry") $\sigma \colon L \cong K$ such that $\sigma X \cdot \sigma Y = X \cdot Y$ for all X and Y in L. The following version of Witt's theorem holds (92:3, [4]): *if* L *and* K *are isometric and if* $L = L' \perp L''$ *and* $K = K' \perp K''$ *with* L' *and* K' *isometric, then* L'' *and* K'' *are isometric.*

The orthogonal group $O(L)$ consists of all isometries of L with itself. If A is a pure vector of L, then the linear mapping τ_A which carries A into $-A$ and is identity on $(\mathfrak{o}A)^\circ$ is clearly in $O(L)$; it is called the *symmetry with respect to* $(\mathfrak{o}A)^\circ$. If A is pure of type i, τ_A is called a symmetry *of type* i. As usual

$$\tau_A X = X - \frac{2(A \cdot X)}{A^2} A \qquad (X \in L).$$

The Cartan-Dieudonné theorem holds for lattices in the following form:

If $\dim L = n$, *every element of* $O(L)$ *can be expressed as a product of*

$\leqq 2n-1$ *symmetries; in addition the number of symmetries of type* i *in the product is* $\leqq 2(\dim L_i)$ *where* (1) *is a Jordan splitting.*

For a proof, see 92:4 and 92:6 in [4]; the statement concerning the number of symmetries of type i is implicit in the proof of 92:4. We note that it is possible to assume that the symmetries are arranged in some particular order according to type—for example those of type 1 coming first, those of type 2 coming second, etc. This follows from the formula $\rho\tau_A\rho^{-1}=\tau_{\rho A}$ (where $\rho\in O(L)$) which shows that the conjugate of a symmetry is another symmetry of the same type.

Now $\det\sigma=\pm 1$ for all $\sigma\in O(L)$; in particular the determinant of a symmetry is -1. By definition, the rotation subgroup $O^+(L)$ is the kernel of $\det\colon O(L)\to\{\pm 1\}$. $O^+(L)$ consists of all mappings in $O(L)$ which are a product of an even number of symmetries.

The identity mapping of a lattice L and the mapping $X\to -X$ are both in $O(L)$ and will be denoted by 1_L and -1_L respectively, or simply by 1 and -1 when no confusion is likely.

If $\mathfrak{a}$ is a non-zero ideal of $\mathfrak{o}$, we define

$$L^{\mathfrak{a}}=\{A\in L\colon A\cdot L\subseteq\mathfrak{a}\}. \tag{2}$$

It is easy to see that $(J\perp K)^{\mathfrak{a}}=J^{\mathfrak{a}}\perp K^{\mathfrak{a}}$, and that if M is modular with $\mathfrak{s}=M\cdot M$, then $M^{\mathfrak{a}}=M$ if $\mathfrak{s}\subseteq\mathfrak{a}$, $M^{\mathfrak{a}}=\mathfrak{a}\mathfrak{s}^{-1}M$ if $\mathfrak{a}\subseteq\mathfrak{s}$. From these facts one can easily get a formula for $L^{\mathfrak{a}}$ if (1) is a Jordan splitting.

It follows immediately from (2) that $\sigma L^{\mathfrak{a}}\subseteq L^{\mathfrak{a}}$ if $\sigma\in O(L)$—in fact $\sigma L^{\mathfrak{a}}=L^{\mathfrak{a}}$. We say that $L^{\mathfrak{a}}$ is an *invariant* lattice (*belonging to* L). This is equivalent to $\sigma\mid L^{\mathfrak{a}}\in O(L^{\mathfrak{a}})$ for all $\sigma\in O(L)$. Thus if $\mathfrak{b}$ is another non-zero ideal, $(L^{\mathfrak{a}})^{\mathfrak{b}}$ is also invariant.

§ 2. Reduction Mod $\mathfrak{p}$.

Throughout this section L is a lattice with Jordan splitting given by (1), and $\mathfrak{s}_i=L_i\cdot L_i$.

Fix a prime element π for $\mathfrak{o}$; thus $\mathfrak{p}=\pi\mathfrak{o}$. For each i, $1\leqq i\leqq t$, we define a map $\kappa_i\colon\mathfrak{s}_i\to k$ by putting $\kappa_i(\alpha)$ equal to the image of $\alpha\pi^{-s(i)}$ under the canonical mapping $\mathfrak{o}\to k$, where $\mathfrak{s}_i=\mathfrak{p}^{s(i)}$. Thus if α and $\beta\in\mathfrak{s}_i$, $\kappa(\alpha)=\kappa(\beta)$ if and only if $\alpha\equiv\beta \bmod \mathfrak{p}\mathfrak{s}_i$.

For each i, $1\leqq i\leqq t$, we put

$$\bar{L}_i=L^{\mathfrak{s}_i}/(L^{\mathfrak{s}_i})^{\mathfrak{p}\mathfrak{s}_i}.$$

$\bar{L}_i$ can be regarded in a natural way as a vector space of dimension $\dim L_i$ over k. When we do this, the map $L^{\mathfrak{s}_i}\to\bar{L}_i$ is semi-linear with respect to $\mathfrak{o}\to k$; we denote the image of $X\in L^{\mathfrak{s}_i}$ by $\bar{X}$.

Since $L^{\mathfrak{s}_i}$ and $(L^{\mathfrak{s}_i})^{\mathfrak{p}\mathfrak{s}_i}$ are both invariant lattices, each σ in $O(L)$ induces a map $\chi_i(\sigma)$ on $\bar{L}_i$ by means of

$$\chi_i(\sigma)(\bar{A}) = \overline{\sigma A}, \qquad (A \in L^{\mathfrak{s}_i}).$$

χ_i is a homomorphism of $O(L)$ into the general linear group $GL(\bar{L}_i)$.

We can also make $\bar{L}_i$ into a regular quadratic space by putting

$$\bar{X} \cdot \bar{Y} = \kappa_i(X \cdot Y), \qquad (X, Y \in L^{\mathfrak{s}_i}).$$

Then $\chi_i(\sigma) \in O(\bar{L}_i)$ for all $\sigma \in O(L)$. In particular it is easy to see that $\chi_i(\tau_A) = \tau_{\bar{A}}$ if A is pure of type i, otherwise $\chi_i(\tau_A) = 1$.

We denote by $O'(V)$ the spinor group (see [1]) of the quadratic space V. We define[2]

$$O^{++}(L) = \bigcap_{i=1}^{t} \chi_i^{-1}(O^{+}(\bar{L}_i)),$$

$$O'(L) = \bigcap_{i=1}^{t} \chi_i^{-1}(O'(\bar{L}_i)).$$

We have $O'(L) \subseteq O^{++}(L)$, and both are normal subgroups of $O(L)$. The group $O^{++}(L)$ can also be described as those elements of $O(L)$ which, when expressed as a product of symmetries, have an even number of each type in the product. In particular $O^{++}(L) \subseteq O^{+}(L)$.

§ 3. Principal Results.

Throughout the rest of the paper L is a lattice of dimension n with a Jordan splitting consisting of t modular components. Unless mentioned otherwise L_i will denote the i-th component in an arbitrary Jordan splitting (1). As before $\mathfrak{s}_i = L_i \cdot L_i$ for $i = 1, \cdots, t$. We shall use the notation and conclusions of § 2 freely.

LEMMA 3.1. *Each element of $O'(L)$ can be written as the product of at most $n - 1$ elements of the form $\tau_A \tau_B$ where $A^2 = B^2$. The elements of this form generate $O'(L)$.*

Proof. It is easy to see that all elements of the stated form are in $O'(L)$, so it suffices to prove the first statement.

By the Cartan-Dieudonné theorem and the fact that $O'(L) \subseteq O^{+}(L)$, every element of $O'(L)$ can be written as the product of $2m \leqq 2n - 2$ symmetries. We shall prove, by induction on m, that an element $\sigma \in O'(L)$ which is expressible as a product of $2m$ symmetries, is the product of exactly m elements of the form required by the lemma. Suppose then that $\sigma = \tau_{A_1}\tau_{A_2} \cdots \tau_{A_{2m}}$.

---

[2] It might appear that $O^{++}(L)$ and $O'(L)$ depend on the choice of the prime element π used in defining κ_i since the bilinear form on $\bar{L}_i$ is given in terms of κ_i. But a change in π merely changes the bilinear form by a non-zero constant factor, and so the groups $O(\bar{L}_i)$, $O^{+}(\bar{L}_i)$ and $O'(\bar{L}_i)$ remain the same.

If $m = 0$, it is trivial so assume $m > 0$. Suppose A_1 is pure of type i. Since $O'(L) \subseteq O^{++}(L)$, there are an even number of symmetries τ_{A_j} of type i. Let $k^{\cdot}$ denote the multiplicative group of k and $k^{\cdot 2}$ its subgroup of squares; $k^{\cdot}/k^{\cdot 2}$ is the cyclic group Z_2 of order 2. Recall that $\chi_i(\tau_A)$ is the identity if A is not pure of type i. From these facts it follows readily that there is an index $j > 1$ such that τ_{A_j} is of type i and $\kappa_i(A_1^2)\kappa_i(A_j^2) \in k^2$, since $\chi_i(\sigma) \in O'(\bar{L}_i)$ is equivalent with $\prod_j \kappa_i(A_j^2) \in k^{\cdot 2}$ where the product is taken over those j with A_j pure of type i. Since the conjugate of a symmetry is a symmetry, we may suppose that $j = 2$. Then $A_1^2 A_2^2 \pi^{-2s(i)}$ is congruent to a square unit mod $\mathfrak{p}$, and so it is actually a square unit by the local square theorem. Thus A_1^2 is equal to a square unit times A_2^2. But $\tau_{\epsilon A} = \tau_A$ for any unit ϵ, and we may therefore suppose that $A_1^2 = A_2^2$. Now $\tau_{A_1}\tau_{A_2} \in O'(L)$, whence $\tau_{A_3} \cdots \tau_{A_{2m}} \in O'(L)$. The result follows by induction.

We let $\Omega(L) = [O(L), O(L)]$ be the commutator subgroup of $O(L)$. Since χ_i is a homomorphism,

$$\chi_i(\Omega(L)) \subseteq \Omega(\bar{L}_i) \subseteq O'(\bar{L}_i) \text{ and so } \Omega(L) \subseteq O'(L).$$

THEOREM[3] 3.2. $\Omega(L) = O'(L)$.

Proof. By the preceding lemma it suffices to show that $\tau_A\tau_B \in \Omega(L)$ if $A^2 = B^2$. By Witt's theorem there is a map $\rho \in O(L)$ which carries A into B. Thus

$$\tau_A\tau_B = \tau_A\tau_{\rho A} = \tau_A \rho \tau_A^{-1} \rho^{-1} \in \Omega(L)$$

as required.

THEOREM 3.3 (Factor groups). *Let t be the number of components in a Jordan splitting of L, and let s be the number of these components with dimension $\geqq 2$. Then $O(L)/\Omega(L)$ is the direct product of the cyclic group Z_2 of two elements with itself $s + t$ times. The factors of the normal series*

$$O(L) \supset O^{+}(L) \supseteq O^{++}(L) \supseteq \Omega(L)$$

are given as follows:

---

[3] This theorem can be stated in another way. Namely let χ be the product of the χ_i so that χ maps $O(\bar{L})$ onto the product of the $O(\bar{L}_i)$. Then $\Omega(L)$ is the pre-image of the commutator subgroup of the product. This latter subgroup is of course the product of the $\Omega(\bar{L}_i)$. A similar theorem holds in the symplectic case since the symplectic group is equal to its commutator subgroup—see (V), p. 127 of my paper, loc. cit.

(i) $O(L)/O^{+}(L)$ *is* Z_2.

(ii) $O^{+}(L)/O^{++}(L)$ *is the direct product of* Z_2 *with itself* $t-1$ *times.*

(iii) $O^{++}(L)/\Omega(L)$ *is the direct product of* Z_2 *with itself* s *times.*

Proof. Define a map D of $O(L)$ into the direct product G of Z_2 with itself t times by

$$D(\sigma) = (\det \chi_1(\sigma), \cdots, \det \chi_t(\sigma)).$$

D is clearly a homomorphism. A consideration of symmetries shows that D is surjective. The kernel of D is $O^{++}(L)$ by definition. Thus $O(L)/O^{++}(L) \cong G$. Now (ii) follows immediately since (i) is trivial.

Let Θ denote the map of $O^{++}(L)$ into G given by

$$\Theta(\sigma) = (\theta\chi_1(\sigma), \cdots, \theta\chi_t(\sigma))$$

where the spinor norm for each of the groups $O^{+}(\bar{L}_i)$ is denoted by θ. The spinor norm of $O^{+}(V)$ for any quadratic space V of dimension 1 is trivial, while for any space of dimension $\geqq 2$ over the finite field k, it is a surjective map onto $\dot{k}/\dot{k}^{2} \cong Z_2$; this follows from the fact that $k = \{X^2 : X \in V\}$ if $\dim V \geqq 2$ (62:1, [4]). By constructing suitable maps in $O^{++}(L)$ as products of symmetries, we see easily that the image of Θ is the direct product of Z_2 with itself s times. Since the kernel of Θ is clearly $O'(L) = \Omega(L)$, (iii) follows immediately.

Suppose $\sigma \in O(L)$. Expressing σ as a product of symmetries and applying Theorem 3.2, we see that $\sigma^2 \in \Omega(L)$. Since (i), (ii) and (iii) imply that $[O(L) : \Omega(L)] = 2^{s+t}$, $O(L)/\Omega(L)$ has the form stated in the theorem.

THEOREM 3.4 (Generators). *$O(L)$ is generated by symmetries. $O^{+}(L)$ (resp. $O^{++}(L)$) is generated by all elements of the form $\tau\tau'$ where τ and τ' are symmetries (resp. symmetries of the same type). $\Omega(L)$ is generated by all products $\tau_A\tau_B$ where $A^2 = B^2$. An element of one of these groups can be written as a product of generators of the prescribed kind, and the minimal number of generators required in such a product is $\leqq 2n-1$ in the case of $O(L)$, and is $\leqq n-1$ in the other three cases.*

Proof. The Cartan-Dieudonné theorem proves the result for $O(L)$, and Lemma 3.1 and Theorem 3.2 prove it for $\Omega(L)$. The proof for $O^{+}(L)$ is trivial. So consider $O^{++}(L)$. By the Cartan-Dieudonné theorem and the remark following it, any element of $O^{++}(L)$ can be written as a product of $\leqq 2n-1$ symmetries where the symmetries of type 1 appear first in the product, those of type 2 next, and so on. But the number of symmetries of each type is even, and so $O^{++}(L)$ is generated by elements of the form $\tau\tau'$

where τ and τ' are symmetries of the same type. But any such element is in $O^{++}(L)$, and the proof of Theorem 3.4 is finished.

The quantity dL_i in the statement of the next theorem is the discriminant of the lattice L_i (p. 222, [4]). It is uniquely determined (up to the square of a unit) by i—that is to say, it does not depend on the Jordan splitting chosen (92:2, [4]).

We denote by cen G the center of the group G.

THEOREM 3.5 (Centers). *The center of* $O(L)$ *is* $\{\pm 1\}$. *If* $\dim L \leqq 2$, *the groups* $O^+(L)$, $O^{++}(L)$ *and* $\Omega(L)$ *are commutative. If* $\dim L \geqq 3$, *the centers of the groups* $O^+(L)$, $O^{++}(L)$ *and* $\Omega(L)$ *are either* $\{1\}$ *or* $\{\pm 1\}$, *and are determined precisely as follows*:

(i) $\operatorname{cen} O^+(L) = \{\pm 1\}$ *if and only if* $\dim L$ *is even.*

(ii) $\operatorname{cen} O^{++}(L) = \{\pm 1\}$ *if and only if* $\dim L_i$ *is even for* $i = 1, \cdots, t$.

(iii) $\operatorname{cen} \Omega(L) = \{\pm 1\}$ *if and only if* $\dim L_i$ *is even and* dL_i *is a square, for* $i = 1, \cdots, t$.

Proof. 1) The case $\dim L = 1$ is trivial since $O(L) = \{\pm 1\}$, so we assume from now on that $\dim L \geqq 2$.

2) Suppose $\dim L = 2$. We shall show that $O^+(L)$ is commutative, whence $O^{++}(L)$ and $\Omega(L)$ are also commutative. By a slightly stronger version of the Cartan-Dieudonné theorem (see the foot-note on p. 248, [4]), any element of $O^+(L)$ is the product of two symmetries, of which one of them may be chosen arbitrarily. Let $\sigma \in O^+(L)$. If τ is any symmetry, then $\sigma = \tau\tau'$ for a suitable symmetry τ' and so

$$\tau\sigma\tau^{-1} = \tau\tau\tau'\tau = \sigma^{-1}.$$

Thus if τ and τ'' are arbitrary symmetries, $(\tau\tau'')\sigma(\tau\tau'')^{-1} = \sigma$, and so $O^+(L)$ is commutative.

3) Let us show that if σ in $O(L)$ carries every component $\mathfrak{o}A$ spanned by a pure vector A into itself, then $\sigma = \pm 1$. For then $\sigma A = \epsilon_A A$ where $\epsilon_A = \pm 1$. Now if A and B are two independent pure vectors, $A + \alpha B$ is pure for a suitable non-zero α in $\mathfrak{o}$ (modular component theorem), and so

$$\sigma(A + \alpha B) = \epsilon_{A+\alpha B}(A + \alpha B) = \epsilon_A A + \alpha\epsilon_B B.$$

Thus $\epsilon_A = \epsilon_{A+\alpha B} = \epsilon_B$, and the result follows.

4) Let us show next that the center of $O(L)$ is $\{\pm 1\}$. If $\sigma \in \operatorname{cen} O(L)$, then $\sigma\tau_A\sigma^{-1} = \tau_A = \tau_{\sigma A}$ whence $\sigma A \in \mathfrak{o}A$ for any pure vector A. Thus $\sigma = \pm 1$ by 3).

5) From now on we assume that $\dim L \geqq 3$. We shall show next that the centralizer G of $\Omega(L)$ in $O(L)$ is $\{\pm 1\}$. It suffices to show that if $\sigma \in G$, then σ carries $\mathfrak{o}A$ into itself for an arbitrary pure vector A.

Suppose A, B and C are three distinct vectors in an orthogonal basis of L. Choose a non-zero $\beta \in \mathfrak{o}$ so that $\beta^2 B^2 \in A^2\mathfrak{p}$ and also so that $A + \beta B$ is pure of the same type as A. Then $\rho = \tau_A \tau_{A+\beta B} \neq 1$ is in $O'(L) = \Omega(L)$, carries $P = \mathfrak{o}A \perp \mathfrak{o}B$ into itself and is identity on P°. Moreover $\rho \mid P$ moves every non-zero vector of P; this can be seen most easily by imbedding P in a two-dimensional quadratic space V over the field of quotients of $\mathfrak{o}$, whose bilinear form extends that on P; then the canonical extension of $\rho \mid P$ to V is in $O^+(V)$, and it follows easily that it moves every non-zero vector of V. Thus P° is characterized as the set of vectors of L stationary under ρ. Since $\sigma\rho\sigma^{-1} = \rho$, the vectors of σP° are stationary under ρ, whence $\sigma P^\circ = P^\circ$, whence $\sigma P = P$. Similarly if $Q = \mathfrak{o}A \perp \mathfrak{o}C$, $\sigma Q = Q$, and so

$$\sigma(\mathfrak{o}A) = \sigma(Q \cap P) = (\sigma Q) \cap (\sigma P) = \mathfrak{o}A$$

as required.

6) By 5), cen $O^+(L) = O^+(L) \cap \{\pm 1\}$, similarly for $O^{++}(L)$ and $\Omega(L)$. Since $\det -1 = 1 \Longleftrightarrow \dim L$ is even, (i) is trivial. Since $\chi_i(-1) = -1$ and $\dim \bar{L}_i = \dim L_i$, (ii) is clear. Now $-1 \in O'(\bar{L}_i) \Longleftrightarrow d\bar{L}_i$ is a square and $\dim \bar{L}_i$ is even (Theorem 5.19, p. 196, [1]). It is straightforward to verify that $d\bar{L}_i$ is a square and $\dim \bar{L}_i$ is even $\Longleftrightarrow dL_i$ is a square and $\dim L_i$ is even, and so (iii) follows. This finishes the proof of Theorem 3.5.

THEOREM 3.6 (Commutator Subgroup). *$\Omega(L)$ is generated by all elements of the form $(\tau\tau')^2$ where τ and τ' are symmetries of the same type. If $\dim L \leqq 2$, the commutator subgroup $[O^+(L), O^+(L)] = 1$, while if $\dim L \geqq 3$, $[O^+(L), O^+(L)] = \Omega(L)$.*

Proof. 1) By Theorem 3.4 $\Omega(L)$ is generated by all maps of the form $\tau_A\tau_B$ with $A^2 = B^2$. To prove the first statement of the theorem it suffices to express any such product in the form $(\tau\tau')^2$ mentioned in the theorem, since all mappings in this latter form are clearly in $\Omega(L)$. By Witt's theorem we can find $\rho \in O(L)$ such that $\rho A = B$, whence $\tau_B = \tau_{\rho A}$. For suitable choice of $\epsilon = \pm 1$, $\rho A + \epsilon A$ is pure of the same type as A. Put $\tau' = \tau_{\rho A + \epsilon A}$. Since $\tau' A = -\epsilon\rho A$, we have $\tau_{\rho A} = \tau_{\tau' A} = \tau' \tau_A \tau'$ and so $\tau_A \tau_B = (\tau_A \tau')^2$ as required.

2) Let $\Omega^+(L) = [O^+(L), O^+(L)]$. If $\dim L \leqq 2$, $\Omega^+(L) = 1$ by Theorem 3.5, so we may assume henceforth that $\dim L \geqq 3$.

To finish the proof, it suffices to show that any element of the form

$(\tau_A\tau_B)^2$, where A and B are not necessarily of the same type, is in $\Omega^+(L)$. Suppose for a moment that $P = \mathfrak{o}A + \mathfrak{o}B$ is not a two-dimensional orthogonal component ("plane") of L. Then it is possible to find a third pure vector C such that $\mathfrak{o}A + \mathfrak{o}C$ and $\mathfrak{o}B + \mathfrak{o}C$ are both planes of L. We have the formula $(\tau_A\tau_B)^2 = (\tau_A\tau_C)^2\rho(\tau_C\tau_B)^2$ where $\rho = (\tau_C\tau_A)(\tau_C\tau_B)(\tau_C\tau_A)^{-1}(\tau_C\tau_B)^{-1} \in \Omega^+(L)$. Therefore we may assume that P is a plane of L.

Let D be a pure vector in P° and put $Q = P \perp \mathfrak{o}D$. Then $L = Q \perp Q^\circ$. Define σ_A (resp. σ_B) to be the mapping which is $-\tau_A$ (resp. $-\tau_B$) on Q and the identity on Q°. Clearly $\sigma_A{}^2 = \sigma_B{}^2 = 1$, $\tau_A\tau_B = \sigma_A\sigma_B$, and σ_A and σ_B are in $O^+(L)$. Then

$$(\tau_A\tau_B)^2 = (\sigma_A\sigma_B)^2 = \sigma_A(\sigma_B\sigma_A)\sigma_A{}^{-1}(\sigma_B\sigma_A)^{-1} \in \Omega^+(L)$$

and the proof of Theorem 3.6 is finished.

UNIVERSITY OF NOTRE DAME.

REFERENCES.

[1] E. Artin, *Geometric Algebra,* Interscience Tracts, New York, 1957.
[2] J. Dieudonné, *La Géométrie des Groupes Classiques,* Seconde Édition, Springer-Verlag, Berlin, 1963.
[3] W. Klingenberg, "Orthogonale Gruppen über lokalen Ringen," *American Journal of Mathematics,* vol. 83 (1961), pp. 281-320.
[4] O. T. O'Meara, *Introduction to Quadratic Forms,* Springer-Verlag, Berlin, 1963.

DERIVATIONS OF OPERATOR GROUP ALGEBRAS.

By Richard V. Kadison and John R. Ringrose.*

1. Introduction. In [2, Theorem 9] Kaplansky proves that each derivation of a type I von Neumann algebra is inner. Establishing a conjecture of Kaplansky, Sakai shows [5] that each derivation of a C^*-algebra is bounded. Using these results Miles [3] notes that each derivation of a C^*-algebra is induced by an operator in the weak closure of some faithful representation of the algebra (a direct sum of irreducible representations from each equivalence class). Again using [2] and [5], it is shown in [1, Theorem 4] that each derivation of a concretely represented C^*-algebra is spatial (i.e., has the form $A \rightarrow BA - AB$ for some bounded operator B). It is also shown in [1, Theorem 7] that each derivation of a hyperfinite von Neumann algebra is inner; that under various assumptions on B the derivation is inner and that the question of whether all derivations of a semifinite von Neumann algebra are inner is equivalent to the question of whether all derivations of a finite von Neumann algebra are inner.

The main result of this paper is:

Theorem 1.1. *Each derivation of the von Neumann algebra generated by the regular representation of a discrete group is inner.*

This result coupled with those of [1] makes it seem very likely that all derivations of von Neumann algebras are inner. It implies, in particular that certain non-hyperfinite factors [4, Lemma 6.3.1] have only inner derivations. We establish it in a sequence of seven lemmas in Section 2, deferring to the last section a Tauberian result which is at the heart of the proof. This Tauberian result, as we need it for the derivation theorem, states that a function on a space acted upon by a transitive permutation group which differs from each of its transforms under the group by a function from a fixed ball in l_2 is itself an l_2 function plus a constant. In the final section, we prove the more general form of this fact corresponding to the action of arbitrary permutation groups. Its application is to the action of a group on a right coset homogeneous space by right translation.

Received April 2, 1965.

* Research conducted with the support of ONR contract NR 043-325 and NSF GP 4059.

We review some facts about operator group algebras and establish some notation. Let G be a (discrete) group, L_a the unitary operator on $l_2(G)$ defined by $L_af(g)=f(a^{-1}g)$ and R_x the unitary operator defined by $R_af(g)=f(ga)$. The mappings $a\rightarrow L_a$ and $a\rightarrow R_a$ are unitary representations of G, the left and right regular representations of G, respectively. The weak closures of finite linear combinations of L_a's and R_a's are von Neumann algebras, the left and right von Neumann group algebras, $\mathcal{L}$ and $\mathcal{R}$, respectively. From [4, Lemma 5.3.4], $\mathcal{R}=\mathcal{L}'$ (the commutant of $\mathcal{L}$).

We shall show that each derivation of $\mathcal{L}$ (equivalently, from [1, Lemma 5], of $\mathcal{R}$) is inner. To this end, we consider the basis $\{x_a\colon a \text{ in } G\}$ for $l_2(G)$ defined by $x_g(h)=\delta_{h,g}$, and determine the properties which the matrix representation of an operator on $l_2(G)$ must have in order that it lie in $\mathcal{L}$, in $\mathcal{R}$, in $\mathcal{L}+\mathcal{R}$ or in $D(\mathcal{L})$ $(=D(\mathcal{R}))$, the set of bounded operators on $l_2(G)$ which induce derivations of $\mathcal{L}$. From [1, Theorem 4], we need establish only that $D(\mathcal{L})=\mathcal{L}+\mathcal{R}$ in order to prove that all derivations of $\mathcal{L}$ (and $\mathcal{R}$) are inner. Assume that A lies in $D(\mathcal{L})$.

For T a bounded operator on $l_2(G)$, let $T(a,b)=(Tx_b,x_a)$ and note that $L_c{}^*TL_c(a,b)=T(ca,cb)$ while $R_c{}^*TR_c(a,b)=T(ac^{-1},bc^{-1})$. Thus T lies in $\mathcal{L}$ if and only if $T(ac,bc)=T(a,b)$ and T lies in $\mathcal{R}$ if and only if $T(ca,cb)=T(a,b)$, for each a, b, c in G (this under the assumption that $(T(a,b))$ is the matrix of a *bounded* operator T). It follows from these considerations that $\mathcal{L}\cap\mathcal{R}$ consists of scalar multiples of the identity operator I (we say that $\mathcal{L}$ and $\mathcal{R}$ are factors) if and only if for each a in G other than the identity e the set (a) of conjugates of a is infinite. Since $T-L_c{}^*TL_c=L_c{}^*(L_cT-TL_c)$, T lies in $D(\mathcal{L})$ if and only if $T-L_c{}^*TL_c$ lies in $\mathcal{L}$ for each c in G (equivalently, if and only if $T-R_c{}^*TR_c$ lies in $\mathcal{R}$ for each c in G). Thus T lies in $D(\mathcal{L})$ if and only if

$$T(a,b)-T(ca,cb)=T(ag,bg)-T(cag,cbg) \tag{1}$$

for each a, b, c and g in G. In words rather than formulas, T lies in $D(\mathcal{L})$, if and only if the difference of two left translates of a matrix coefficient of T is right invariant.

We are grateful to H. Sah for discussions of group theoretic constructs which led us to consider the Tauberian result (Theorem 3.8) in the setting of a permutation group acting on a set.

2. The main result. If we knew that $A=B+C$ with B in $\mathcal{L}$ and C in $\mathcal{R}$, then $A(g,ag)=B(g,ag)+C(g,ag)=B(e,a)+C(e,g^{-1}ag)$; so that as g ranges through distinct representatives of the cosets Z_ag in G/Z_a,

$C(e, g^{-1}ag)$ ranges through distinct matrix coefficients in the e row of $(C(a, b))$. In particular we should have that A_a is a function on G/Z_a, where Z_a is the centralizer of a and $A_a(Z_a g) = A(g, ag)$. Since $(C(a, b))$ is to be the matrix of a bounded operator, each of its rows is in $l_2(G)$, and we should have A_a tends to the limit $B(e, a)$ at ∞ on G/Z_a if (a) is infinite. If $(A(a, b))$ satisfies the condition:

(2) $\quad A(ca, cb) = A(a, b)$, for each c in G when (ba^{-1}) is finite

then $B(e, ghg^{-1}) = A(g, gh) - C(e, h) = A(e, h) - C(e, h) = B(e, h)$ for each g in G when (h) is finite, i.e., $g \rightarrow B(e, g)$ is constant on (h) and it turns out that we may choose 0 as this constant value. The sequence of lemmas which follows establishes that A_a is a function on G/Z_a which has a limit at ∞, that A can be replaced by an operator whose matrix satisfies (2) and that, after this replacement is made, $B(e, a)$ as described is the e row of a matrix corresponding to an operator B in $\mathcal{L}$ such that $A - B$ $(= C)$ lies in $\mathcal{R}$.

We adopt the usual terminology that the point added in forming the one point compactification of a locally compact space is the point ∞ and that the behavior of a function on the locally compact space at ∞ is the behavior of the function on the one-point compactification in the neighborhood of ∞. In particular we may speak of $\overline{\lim} f$ and $\underline{\lim} f$ at ∞, as well as $\lim f$ at ∞ (if this limit exists). We make use of Lemma 3.7, in the following lemma and Lemma 3.1 in the proof of Lemma 2.7, deferring until the final section the proofs of Lemmas 3.1 and 3.7.

LEMMA 2.1. *A_a is single-valued and has a limit at ∞ on G/Z_a.*

Proof. We wish to show that $A(bg, abg) = A(g, ag)$, when $ba = ab$. Since $A(bg, abg) = A(bg, bag)$, we have, from (1)

$$\begin{aligned}\alpha = A(g, ag) - A(bg, bag) &= A(e, a) - A(b, ba) = A(b, ab) - A(b^2, bab)\\ &= A(b, ba) - A(b^2, b^2a) = \cdots = A(b^{n-1}, b^{n-1}a) - A(b^n, b^n a).\end{aligned}$$

Thus $n\alpha = A(e, a) - A(b^n, b^n a)$ and $n|\alpha| \leqq 2\|A\|$ for each n. Hence $\alpha = 0$.

Note that

$$\begin{aligned}\sum_{G/Z_a} |A_a(Z_a g) - A_a(Z_a gh)|^2 &= \sum_n |A(g_n, ag_n) - A(g_n h, ag_n h)|^2\\ &= \sum_n |A(e, g_n^{-1}ag_n) - A(h, g_n^{-1}ag_n h)|^2 \leqq 4\|A\|^2\end{aligned}$$

where $\{g_n\}$ is a complete set of representatives of the cosets in G/Z_a. This inequality results from the fact that $\{A(e, g_n^{-1}ag_n)\}$ and $\{A(h, g_n^{-1}ag_nh)\}$ are distinct elements of the e and h rows for the matrix for A and the known relation of the l_2-norm of such rows to the bound of the operator.

Since G acts as a transitive permutation group on G/Z_a by right multiplication, Lemma 3.7 applies and A_a has a limit at ∞.

LEMMA 2.2. *If* $(A(a,b))$ *satisfies* (2) *and* $C(a,b) = A(a,b) - B(e, ba^{-1})$, *where* $B(e,h) = 0$ *if* (h) *is finite and* $B(e,h) = \lim_\infty A_h$ *if* (h) *is infinite, then* $C(ca, cb) = C(a,b)$, *for each* a, b *and* c *in* G.

Proof. If (ba^{-1}) is finite,

$$C(ca, cb) = A(ca, cb) - B(e, cba^{-1}c^{-1})$$
$$= A(ca, cb) = A(a,b) = A(a,b) - B(e, ba^{-1}) = C(a,b),$$

from (2).

Suppose that (ba^{-1}) is infinite. We wish to show:

(3) $A(ca, cb) - A(a,b) = B(e, cba^{-1}c^{-1}) - B(e, ba^{-1})$, for each c in G.

If $\{g_n\}$ is a complete set of representatives for the cosets in $G/Z_{ba^{-1}}$, then $\{cg_nc^{-1}\}$, hence $\{cg_n\}$, are complete sets of representatives for the cosets in $G/Z_{cba^{-1}c^{-1}}$. Thus, with positive ϵ assigned, for suitable n,

$$\epsilon > | B(e, ba^{-1}) - A(g_n, ba^{-1}g_n) | + | A(cg_n, cba^{-1}g_n) - B(e, cba^{-1}c^{-1}) |$$
$$\geqq | B(e, ba^{-1}) - B(e, cba^{-1}c^{-1}) - [A(g_ng_n^{-1}a, ba^{-1}g_ng_n^{-1}a)$$
$$- A(cg_ng_n^{-1}a, cba^{-1}g_ng_n^{-1}a)] |,$$

from which (3) follows.

LEMMA 2.3. *If* $B_a(Z_ag) = B(e, g^{-1}ag)$ *and* $C_a(Z_ag) = C(e, g^{-1}ag)$, *then with* (a) *infinite* B_a *and* C_a *vanish at* ∞ *on* G/Z_a.

Proof. From Lemma 2.2, we have $A(g, ag) = B(e, a) + C(e, g^{-1}ag)$. Since $A_a(Z_ag) = A(g, ag)$ and $B(e,a)$ is $\lim_\infty A_a$ on G/Z_a, C_a vanishes at ∞ on G/Z_a. But A'_a defined on G/Z_a by

$A'_a(Z_ag) = A(e, g^{-1}ag)$ $(= B(e, g^{-1}ag) + C(e, g^{-1}ag) = B_a(Z_ag) + C_a(Z_ag))$ vanishes at ∞, since the e row of $(A(a,b))$ lies in $l_2(G)$. Since C_a vanishes at ∞ on G/Z_a, B_a $(= A'_a - C_a)$ does.

LEMMA 2.4. *If* $bhg^{-1}b^{-1} = ahg^{-1}a^{-1}$ *then* $A(bg, bh) = A(ag, ah)$.

Proof. Since $(ag)^{-1}(ahg^{-1}a^{-1})ag = (bg)^{-1}(bhg^{-1}b^{-1})bg$ and $ahg^{-1}a^{-1} = bhg^{-1}b^{-1}$, we have $A(ag, ah) = A(ag, (ahg^{-1}a^{-1})ag) = A(bg, (bhg^{-1}b^{-1})bg)$
and only if that induced by A is. If ba^{-1} lies in S, then $A^{|S}(ca, cb) = A^{|S}(a,b)$;
$= A(bg, bh)$, from Lemma 2.1.

For each finite (h), we choose a set $a_1, \cdots, a_n$ of elements in G which give rise to all (distinct) permutations of (h) induced by inner automorphisms of G; and we denote by $A^{|h}$ the operator $\frac{1}{n}(L_{a_1}{}^*AL_{a_1} + \cdots + L_{a_n}{}^*AL_{a_n})$ (we write $A^{|h_1|h_2}$ for $(A^{|h_1})^{|h_2}$).

LEMMA 2.5. *If* (h_j) *is finite for* $j = 1, \cdots, m$, then $\| A^{|h_1\cdots|h_m} \| \leqq \| A \|$, $A^{|h_1\cdots|h_m}$ *induces the same derivation of* $\mathfrak{R}$ *as* A *does and*

$$A^{|h_1\cdots|h_m}(ca, cb) = A^{|h_1\cdots|h_m}(a, b)$$

for each c *in* G *when* $ba^{-1} = h_j$ *for some* j.

Proof. Since each L_g is a unitary operator in $\mathcal{L}$, if T induces a derivation of $\mathcal{L}$, $L_g{}^*TL_g$ induces the same derivation of $\mathfrak{R}$ as T does (cf. [1, Lemma 5]), and $\| L_g{}^*TL_g \| = \| T \|$.

The first and second assertions follow from this and the definition of $T^{|h}$.

With $T^{|h_j} = \frac{1}{n}(L_{a_1}{}^*TL_{a_1} + \cdots + L_{a_n}{}^*TL_{a_n})$, we have

$$T^{|h_j}(ca, cb) = \frac{1}{n}\sum_j T(a_jca, a_jcb) = \frac{1}{n}\sum_j T(a_ja, a_jb) = T^{|h_j}(a, b),$$

from Lemma 2.4, since $a_{j'}ba^{-1}a_{j'}{}^{-1} = a_jcba^{-1}c^{-1}a_j{}^{-1}$ for some permutation $j \to j'$ of $\{1, \cdots, n\}$, by choice of $a_1, \cdots, a_n$. Hence

$$L_g{}^*T^{|h_j}L_g(ca, cb) = T^{|h_j}(gca, gcb) = T^{|h_j}(a, b)$$

for each g in G; so that

$$T^{|h_j|h_{j+1}}(ca, cb) = T^{|h_j|h_{j+1}}(a, b) \text{ and } T^{|h_j\cdots|h_m}(ca, cb) = T^{|h_j\cdots|h_m}(a, b).$$

Replacing T by $A^{|h_1\cdots|h_{j-1}}$ completes the proof.

For each finite subset S of G_0, the subgroup of G consisting of elements h with (h) finite, we assign a linear order $h_1, h_2, \cdots, h_m$ to its members and denote $A^{|h_1\cdots|h_m}$ by $A^{|S}$.

LEMMA 2.6. *Relative to the family* $\mathfrak{S}$ *of finite subsets of* G_0 *directed by inclusion, the net* $\{A^{|S}\}$ *has a (cofinal) subnet weak-operator convergent to an operator* A_0 *satisfying* $A_0(ca, cb) = A_0(a, b)$ *for each* c *in* G *and each* a, b *in* G *such that* (ba^{-1}) *is finite.* A_0 *induces a derivation of* $\mathcal{L}$ *which is inner if and only if that induced by* A *is.*

Proof. Since $\| A^{|S} \| \leqq \| A \|$, the ball of radius $\| A \|$ is weak-operator compact, and $A^{|S}$ induces the same derivation of $\mathcal{R}$ as A does, the net $\{A^{|S}\}$ has a subnet convergent to A_0, and A_0 induces the same derivation of $\mathcal{R}$ as A does. From [1, Lemma 5], A_0 induces a derivation of $\mathcal{L}$ which is inner if and only if that induced by A is. If ba^{-1} lies in S, then $A^{|S}(ca, cb) = A^{|S}(a, b)$; so that $A_0(ca, cb) = A_0(a, b)$, for each c in G, from Lemma 2.5 and the fact that A_0 is the limit of a (cofinal) subnet of $\{A^{|S}\}$.

We may assume, henceforth, that A satisfies (2) (replacing A by A_0).

LEMMA 2.7. *Taking $B(a,b)$ as $B(e,ba^{-1})$, $(B(a,b))$ is the matrix of a (bounded) operator B in $\mathcal{L}$.*

Proof. Let $\mathcal{F}$ be the family of finite subsets of $\bigcup_{(a)\text{ infinite}} G/Z_a$ directed by inclusion. From Lemma 3.1, we can choose an element a_S^{-1} in G, for each S in $\mathcal{F}$, which lies in none of the cosets in S. The net $\{A - L_{a_S}{}^*AL_{a_S}\}$ lies in the weak-operator compact ball of radius $2\,\|A\|$ in $\mathcal{L}$ and therefore has a (cofinal) subnet weak-operator convergent to some T in this ball.

We show that $T(a,b) = T(e,ba^{-1}) = B(a,b) = B(e,ba^{-1})$. Note that

$$
\begin{aligned}
&(A - L_{a_S}{}^*AL_{a_S})(e,h)\\
&\qquad = A(e,h) - A(a_S, a_S h) = B(e,h) - B(e, a_S h a_S^{-1})\\
&\qquad = B(e,h) - B_h(Z_h a_S^{-1}),
\end{aligned}
$$

by Lemma 2.2. Now B_h tends to 0 at ∞ on G/Z_h, for $(A(a,b))$ satisfies (2) so that $B_h(Z_h a^{-1}) = B(e, aha^{-1}) = 0$, with (h) finite; and with (h) infinite Lemma 2.3 applies. Thus $B(e,h) - B_h(Z_h a_S^{-1})$ tends to $B(e,h)$ over $\mathcal{F}$, for if S contains a suitable finite subset of G/Z_h, $|B_h(Z_h a_S^{-1})|$ is small since $Z_h a_S^{-1}$ is not in this subset by choice of a_S. Hence $T(e,h) = B(e,h)$ for each h; so that $(B(a,b))$ is the matrix for T.

Proof of Theorem 1.1. Since $C(a,b) = A(a,b) - B(a,b)$, $(C(a,b))$ is the matrix of a bounded operator C. Since $C(ca,cb) = C(a,b)$ (Lemma 2.2), C lies in $\mathcal{R}$. Since $A = B + C$ with B in $\mathcal{L}$ and C in $\mathcal{R}$, the derivation of $\mathcal{L}$ induced by A is inner.

3. A Tauberian result.

We now prove the results, concerning groups and in particular permutation groups, which were assumed in the previous section. The first of these is the following.

LEMMA 3.1. *A group G cannot be expressed as the union of a finite number of right (left) cosets of subgroups with infinite index.*

Proof. Suppose the contrary, and that $H_1, \cdots, H_n$ is a minimal set of subgroups of G, each with infinite index in G, such that G can be expressed as the union of a finite number of (right) cosets of $H_1, \cdots, H_n$. Let $H_1g_1, \cdots, H_1g_k$ be the cosets of H_1 which appear in some such expression for G, and let H_1g be a coset distinct from each H_1g_j $(j=1, \cdots, k)$. Then

$$H_1g \subseteq G - H_1g_1 \cup \cdots \cup H_1g_k,$$

and since the right hand side is contained in a finite union of cosets of $H_2, \cdots, H_n$, the same is true of H_1g and therefore, also, of each H_1g_j $(j=1, \cdots, k)$. It follows that G can be expressed as a finite union of cosets of $H_2, \cdots, H_n$, contrary to the minimal nature of the set $H_1, \cdots, H_n$. This contradiction proves the lemma.

COROLLARY 3.2. *Let* Π *be a group of permutations of a set* X, *and let* S *be a finite subset of* X *such that, for each* x *in* S, *the orbit* $\{\pi(x) : \pi \in \Pi\}$ *is infinite. Then there exists* π *in* Π *such that the sets* S *and* $\pi(S)$ *are mutually disjoint.*

Proof. Given x and y in X, let $H_{x,y} = \{\pi : \pi \in \Pi \text{ and } \pi(y) = x\}$. Then $H_{x,x}$ is a subgroup of Π, and $H_{x,y}$ is either empty or a right coset of $H_{x,x}$. If $x \in S$ then $H_{x,x}$ has infinite index in Π, since $H_{x,y}$ runs through an infinity of distinct cosets of $H_{x,x}$ as y runs through the (infinite) orbit of x. By Lemma 3.1 there exists π in Π such that $\pi \notin H_{x,y} (x, y \in S)$; and the sets S and $\pi(S)$ are then mutually disjoint.

Before proving the second result which was assumed in Section 2 (Lemma 3.7), we require the following definition and two auxiliary lemmas.

Definition 3.3. Let Π be a group of permutations of an infinite set X, and let Y and Z be mutually disjoint infinite subsets of X. We say that Y "penetrates" Z, and write "$Y \leftrightarrow Z$" if, given any positive integer n, there exists π in Π such that $\pi(Y) \cap Z$ contains at least n members. If Y does not penetrate Z, we write "$Y \not\leftrightarrow Z$."

Remark 3.4. Let Y and Z be mutually disjoint infinite subsets of X, and let R be a finite subset of X. Since, for each π in Π,

$$\operatorname{card}\{\pi(Y) \cap Z\} = \operatorname{card}\{Y \cap \pi^{-1}(Z)\},$$

$$\operatorname{card}\{\pi(Y) \cap Z\} \geqq \operatorname{card}\{\pi(Y - R) \cap (Z \cup R)\} - \operatorname{card}\{R\}$$

(where "card" denotes the cardinal of the set in question), it is apparent that $Y \leftrightarrow Z$ if and only if $Z \leftrightarrow Y$, and that $Y \leftrightarrow Z$ if $Y - R \leftrightarrow Z \cup R$. Furthermore, if $Y_1, \cdots, Y_p, Z_1, \cdots, Z_q$ are pairwise disjoint infinite subsets of X, then

$$Y_1 \cup Y_2 \cup \cdots \cup Y_p \leftrightarrow Z_1 \cup Z_2 \cup \cdots \cup Z_q$$

if and only if there exist j and k such that $Y_j \leftrightarrow Z_k$; for

$$\operatorname{card}\{\pi(\bigcup_j Y_j) \cap \bigcup_k Z_k)\} = \sum_{j,k} \operatorname{card}\{\pi(Y_j) \cap Z_k\},$$

and the right hand side of this equation is unbounded, as π runs through Π, if and only if at least one of the summands $\operatorname{card}\{\pi(Y_j) \cap Z_k\}$ is unbounded.

LEMMA 3.5. *Let* Π *be a transitive group of permutations of an infinite set* X. *Let* $Y_1, \cdots, Y_n$ *and* $Z_1, \cdots, Z_n$ *be infinite subsets of* X *such that*

(i) $Y_1 \subseteq Y_2 \subseteq \cdots \subseteq Y_n$, $Z_1 \subseteq Z_2 \subseteq \cdots \subseteq Z_n$,

(ii) $Y_m \not\leftrightarrow X - Y_{m+1}$ *and* $Z_m \not\leftrightarrow X - Z_{m+1}$ $(m = 1, \cdots, n-1)$,

(iii) Y_n *and* Z_n *are mutually disjoint.*

Then there exists π *in* Π *such that* $\pi(Y_n) \cap Z_n$ *contains at least* n *elements.*

Proof. We shall prove by induction on m that, for $m = 1, \cdots, n$, the following statement $[m]$ holds: $[m]$ there exist a sequence $(Q_1, Q_2, \cdots)$ of pairwise disjoint subsets of Y_m, each containing just m elements, and a sequence $(\rho_1, \rho_2, \cdots)$ of members of Π, such that $(\rho_1(Q_1), \rho_2(Q_2), \cdots)$ is a sequence of pairwise disjoint subsets of Z_m. It is clear that the statement $[n]$ implies the conclusion of the lemma.

Since Y_1 and Z_1 are infinite sets, the assertion $[1]$ follows at once from the transitivity of Π. Suppose that $[m]$ has been established for some m satisfying $1 \leqq m < n$, and that Q_j and ρ_j $(j = 1, 2, \cdots)$ have been chosen in accordance with $[m]$. We shall prove $[m+1]$ in the following form: $[m+1]$ there exist a sequence $(R_1, R_2, \cdots)$ of pairwise disjoint subsets of Y_{m+1}, each containing just $m+1$ elements, and a sequence $(\sigma_1, \sigma_2, \cdots)$ of members of Π, such that $(\sigma_1(R_1), \sigma_2(R_2), \cdots)$ is a sequence of pairwise disjoint subsets of Z_{m+1}.

Suppose that k is a positive integer, and that suitable σ_j and R_j have been chosen for j (if any) such that $1 \leqq j < k$. We shall prove the existence of suitable σ_k and R_k.

We begin by noting that there exists x in Y_m such that, for an infinity of values of r, $\rho_r^{-1}(x) \in Y_{m+1} - R$, where R is the finite set $R_1 \cup R_2 \cup \cdots \cup R_{k-1}$. For suppose that no such x exists. Then, given any finite subset Q of $Y_m - R$, we have $\rho_r^{-1}(Q) \subseteq (X - Y_{m+1}) \cup R$ for all sufficiently large r. It follows that $Y_m - R \leftrightarrow (X - Y_{m+1}) \cup R$, and hence that $Y_m \leftrightarrow X - Y_{m+1}$, contrary to hypothesis. This proves the existence of x in Y_m with the stated property. If we replace (ρ_r), and correspondingly (Q_r), by suitable subsequences, we may suppose that $\rho_r^{-1}(x) \in Y_{m+1} - R$ $(r = 1, 2, \cdots)$. Since $\rho_r(Q_r) \subseteq Z_m \subseteq X - Y_m$, while $\rho_r(\rho_r^{-1}(x)) = x \in Y_m$, it follows that $\rho_r^{-1}(x) \notin Q_r$. By recalling the definition of R, we now obtain

$$(4) \qquad \rho_r^{-1}(x) \in Y_{m+1} - R_1 \cup R_2 \cup \cdots \cup R_{k-1}, \qquad \rho_r^{-1}(x) \notin Q_r.$$

Since Π acts transitively on X, there exists π in Π such that

$$(5) \qquad \pi(x) \in Z_{m+1} - \sigma_1(R_1) \cup \sigma_2(R_2) \cup \cdots \cup \sigma_{k-1}(R_{k-1}).$$

We assert that only a finite number of the sets $(\pi\rho_1)(Q_1), (\pi\rho_2)(Q_2), \cdots$ meet $(X - Z_{m+1}) \cup R'$, where

$$R' = \sigma_1(R_1) \cup \sigma_2(R_2) \cup \cdots \cup \sigma_{k-1}(R_{k-1}).$$

For suppose this assertion is false. Then since $\rho_1(Q_1), \rho_2(Q_2), \cdots$ are pairwise disjoint subsets of Z_m, and an infinity of their images $(\pi\rho_1)(Q_1), (\pi\rho_2)(Q_2), \cdots$ under π meet $(X - Z_{m+1}) \cup R'$, it follows that π maps an infinite subset of Z_m into $(X - Z_{m+1}) \cup R'$. Since R' is finite, π maps an infinite subset of Z_m into $X - Z_{m+1}$, and so $Z_m \leftrightarrow X - Z_{m+1}$, contrary to hypothesis. This proves the assertion made in the first sentence of this paragraph. Furthermore, since $Q_1, Q_2, \cdots$ are pairwise disjoint, only a finite number of them meet the finite set $R_1 \cup R_2 \cup \cdots \cup R_{k-1}$. It follows that, by avoiding a finite set of values, we may choose r such that

$$(6) \qquad Q_r \subseteq Y_m - R_1 \cup R_2 \cup \cdots \cup R_{k-1},$$

$$(7) \qquad (\pi\rho_r)(Q_r) \subseteq Z_{m+1} - \sigma_1(R_1) \cup \sigma_2(R_2) \cup \cdots \cup \sigma_{k-1}(R_{k-1}).$$

If we now define

$$R_k = Q_r \cup \{\rho_r^{-1}(x)\}, \sigma_k = \pi\rho_r$$

then it follows at once from (4), $\cdots$, (7) that R_k consists of just $m + 1$ elements of Y_{m+1} and does not meet any of the sets $R_1, \cdots, R_{k-1}$, while $\sigma_k(R_k)$ is a subset of Z_{m+1} which does not meet any of the set $\sigma_1(R_1), \cdots, \sigma_{k-1}(R_{k-1})$.

This inductive construction for R_k and σ_k proves the statement $[m + 1]$, and so completes the proof of Lemma 3.5.

LEMMA 3.6. *Let Π be a transitive group of permutations of an infinite set X. Then*

(i) *if $X = Y \cup Z$, where Y and Z are mutually disjoint infinite subsets of X, then $Y \leftrightarrow Z$;*

(ii) *if, for each $n = 2, 3, \cdots, X$ is expressed in the form $X = \bigcup_{j=1}^{2^n} X_j^{(n)}$,* *where the sets* $X_j^{(n)}$ $(j=1,\cdot\cdot\cdot,2^n)$ *are infinite and pairwise disjoint, and*

$$(8) \qquad X_j^{(n)} = X_{2j-1}^{(n+1)} \cup X_{2j}^{(n+1)} \qquad (j=1,\cdot\cdot\cdot,2^n;\, n=2,3,\cdot\cdot\cdot),$$

then there exist integers n, j, k *such that* $n \geqq 2$, $1 \leqq j < j+2 \leqq k \leqq 2^n$, *and* $X_j^{(n)} \leftrightarrow X_k^{(n)}$.

Proof. (i) Suppose that $Y \not\leftrightarrow Z$. Given any positive integer n, define $Y_m = Y$, $Z_m = Z$ $(m=1,\cdot\cdot\cdot,n)$. The hypotheses of Lemma 3.5 are satisfied, so there exists π in Π such that $\pi(Y) \cap Z$ contains at least n elements. Hence, despite our assumption to the contrary, $Y \leftrightarrow Z$.

(ii) Suppose that there are no integers n, j, k with the stated properties. Given any n $(\geqq 2)$, we define Y_m and Z_m $(m=1,\cdot\cdot\cdot,2^{n-2})$ by

$$(9) \qquad Y_m = \bigcup_{r=1}^{m} X_r^{(n)}, \qquad Z_m = \bigcup_{r=1}^{m} X_{2^n+1-r}^{(n)}.$$

Since

$$X - Y_{m+1} = \bigcup_{r=m+2}^{2^n} X_r^{(n)}, \qquad X - Z_{m+1} = \bigcup_{r=m+2}^{2^n} X_{2^n+1-r}^{(n)},$$

our assumption that $X_j^{(n)} \not\leftrightarrow X_k^{(n)}$ when $k \geqq j+2$ implies that

$$Y_m \not\leftrightarrow X - Y_{m+1} \text{ and } Z_m \not\leftrightarrow X - Z_{m+1}$$

(cf. Remark 3.4). Hence the conditions of Lemma 3.5 are satisfied (with n replaced by 2^{n-2}), so there exists π in Π such that

$$\pi(Y_{2^{n-2}}) \cap Z_{2^{n-2}}$$

contains at least 2^{n-2} elements. Since, by virtue of (8) and (9),

$$Y_{2^{n-2}} = X_1^{(2)}, \qquad Z_{2^{n-2}} = X_4^{(2)},$$

we have shown that, given any integer n $(\geqq 2)$, there exists π in Π such that $\pi(X_1^{(2)}) \cap X_4^{(2)}$ contains at least 2^{n-2} members. Thus $X_1^{(2)} \leftrightarrow X_4^{(2)}$ and, despite our assumption to the contrary, the conclusions of part (ii) of the lemma are satisfied when $n=2$, $j=1$, $k=4$. This contradiction implies the existence of some n, j, k with the desired properties.

The following result will be subsumed in Theorem 3.8, but is needed during its proof.

LEMMA 3.7. *Let* Π *be a group of permutations of a set* X, *let* M *be a positive real number, and let* u *be a complex valued function which is defined on* X *and satisfies*

$$\sum_{x \in X} |u(x) - u(\pi(x))|^2 \leqq M^2 \quad (\pi \in \Pi). \tag{10}$$

Let Z be an infinite orbit. Then $\lim_\infty u$ *exists on* Z.

Proof. Let $u' = u \mid Z$, $\Pi' = \{\pi \mid Z : \pi \in \Pi\}$. Then Π' is a transitive group of permutations of Z, and u' satisfies the same hypotheses relative to Z and Π' as does u relative to X and Π. We have to show that $\lim_\infty u'$ exists on Z. Hence it is sufficient to consider the case in which Π acts transitively on X. Furthermore, we may assume without loss of generality that u is a real valued function.

Suppose therefore that Π is a transitive group of permutations on an infinite set X, and that u is a real valued function which is defined on X and satisfies (10). Given x and y in X, there exists π in Π such that $y = \pi(x)$, and it follows from (10) that $|u(x) - u(y)| \leqq M$. Hence u is bounded on X, and therefore has finite lower and upper limits. We define

$$a = \underline{\lim}_\infty u, \qquad b = \overline{\lim}_\infty u. \tag{11}$$

We shall assume the non-existence of $\lim_\infty u$, so that $a < b$; in due course we shall obtain a contradiction.

We begin by proving the existence of real numbers c and d satisfying $a < c < d < b$ and such that, if

$$Q = \{x : x \in X, u(x) < c\}, \qquad R = \{x : x \in X, u(x) \geqq d\}, \tag{12}$$

then $Q \leftrightarrow R$. We deal separately with two cases.

Case 1. *The closure of the set* $\{u(x) : x \in X\}$ *does not contain the whole of the compact interval* $[a, b]$. In this case we may choose c and d so that $a < c < d < b$ and $u(x) \notin [c, d]$ $(x \in X)$. The sets Q and R defined by (12) are infinite (by (11)) and mutually disjoint, and $Q \cup R = X$. By Lemma 3.6 (i), $Q \leftrightarrow R$.

Case 2. *The closure of* $\{u(x) : x \in X\}$ *contains* $[a, b]$. Given any integer n $(\geqq 2)$, let $d_n = 2^{-n}(b - a)$, and define $X_j^{(n)}$ $(j = 1, \cdots, 2^n)$ by

$$X_j^{(n)} = \{x : x \in X, a + (j-1)d_n \leqq u(x) < a + jd_n\} \quad (j = 2, \cdots, 2^n - 1),$$
$$X_1^{(n)} = \{x : x \in X, u(x) < a + d_n\} \quad \text{and} \quad X_{2^n}^{(n)} = \{x : x \in X, u(x) \geqq b - d_n\}.$$

The hypotheses of Lemma 3.6 (ii) are satisfied, so there exist integers n, j, k such that $n \geqq 2$, $1 \leqq j < j + 2 \leqq k \leqq 2^n$, and $X_j^{(n)} \leftrightarrow X_k^{(n)}$. If we now take $c = a + jd_n$, $d = a + (k-1)d_n$, and define Q and R by (12), then $X_j^{(n)} \subseteq Q$ and $X_k^{(n)} \subseteq R$, so $Q \leftrightarrow R$.

In both cases, we have proved the existence of real numbers c and d

with the stated properties. Suppose now that n is any positive integer. Since $Q \leftrightarrow R$, we may find π in Π and a set S consisting of n members of Q, such that $\pi(S) \subseteq R$. From (10) and (12) it follows that

$$M^2 \geqq \sum_{x \in S} |u(x) - u(\pi(x))|^2 \geqq n(d-c)^2.$$

If n is sufficiently large, we have a contradiction, so the lemma is proved.

THEOREM 3.8. *Let Π be a group of permutations of a set X, let M be a positive real number, and let u be a complex valued function which is defined on X and satisfies*

$$(13) \qquad \sum_{x \in X} |u(x) - u(\pi(x))|^2 \leqq M^2 \quad (\pi \in \Pi).$$

Then u can be expressed (uniquely) in the form $u_1 + u_2$, where u_1 and u_2 are complex valued functions defined on X, u_2 is constant on each orbit and is bounded if u is bounded, while

$$(14) \qquad \sum_{x \in X} |u_1(x)|^2 \leqq \tfrac{1}{2}M^2,$$

and u_1 satisfies the following condition: if Y is the union of a finite family of finite orbits, and $\Pi' = \{\pi \mid Y : \pi \in \Pi\}$, then

$$(15) \qquad \sum_{\pi' \in \Pi'} u_1(\pi'(y)) = 0 \quad (y \in Y).$$

Proof. Let $x \in X$. If the orbit $\Pi(x)$ of x is finite, we define $u_2(x)$ to be the mean value of u on $\Pi(x)$. If $\Pi(x)$ is infinite, we define $u_2(x)$ to be $\lim_\infty u$ on $\Pi(x)$; the existence of the limit follows from Lemma 3.7. The function u_2 defined in this way is constant on each orbit, and is bounded if u is bounded. Let $u_1 = u - u_2$. Then u_1 has mean value zero on each finite orbit, and has limit zero on each infinite orbit. Furthermore, since $u_2(x) = u_2(\pi(x))$ whenever $x \in X$ and $\pi \in \Pi$, it follows that u_1 satisfies (13).

Let Y be the union of a finite family of finite orbits, and let $\Pi' = \{\pi \mid Y : \pi \in \Pi\}$. If $y \in Y$ and $\rho, \sigma \in \Pi'$, then $\rho(y) = \sigma(y)$ if and only if ρ and σ belong to the same left coset of the subgroup $\{\pi' : \pi' \in \Pi', \pi'(y) = y\}$ of Π'; so, if n is the order of this subgroup, and $z \in \Pi(y)$, then there are just n elements π' in Π' such that $\pi'(y) = z$. Since u_1 has mean value zero on $\Pi(y)$,

$$\sum_{\pi' \in \Pi'} u_1(\pi'(y)) = n \sum_{z \in \Pi(y)} u_1(z) = 0.$$

Apart from the (straightforward) question of the uniqueness of u_1 and u_2, it remains only to prove (14). For this, it is sufficient to show that, if

Q is any finite subset of X and ϵ is any positive real number, then there exist complex numbers $a(x)$ and $b(x)$ $(x \in Q)$ such that

$$|a(x)| < \epsilon, \qquad |b(x)| < \epsilon \quad (x \in Q) \tag{16}$$

and

$$\sum_{x \in Q} |u_1(x) - a(x)|^2 + \sum_{x \in Q} |u_1(x) - b(x)|^2 \leqq M^2. \tag{17}$$

Suppose that we are given such Q and ϵ. We define

$$R = \{x \colon x \in Q \text{ and } \Pi(x) \text{ is finite}\}, \ Y = \bigcup_{x \in R} \Pi(x), \tag{18}$$

$$S = \{x \colon x \in Q \text{ and } \Pi(x) \text{ is infinite}\}, \ Z = \bigcup_{x \in S} \Pi(x). \tag{19}$$

Then Y is the union of a finite family of finite orbits. Let $\Pi' = \{\pi \mid Y \colon \pi \in \Pi\}$, let k be the order of Π', and choose $\rho_1, \cdots, \rho_k$ in Π such that $\Pi' = \{\rho_j \mid Y \colon j = 1, \cdots, k\}$.

Since u_1 has limit zero on each infinite orbit, while Z is the union of a finite family of infinite orbits, it follows that u_1 has limit zero on Z. Hence the sets S_1 and S_2, defined by

$$\begin{aligned} S_1 &= S \cup \{x \colon x \in Z \text{ and } |u_1(x)| \geqq \epsilon\}, \\ S_2 &= S_1 \cup \rho_1(S_1) \cup \cdots \cup \rho_k(S_1), \end{aligned} \tag{20}$$

are finite subsets of Z. By Corollary 3.2, there exists π in Π such that the sets S_2 and $\pi(S_2)$ are mutually disjoint. In particular, for each j $(= 1, \cdots, k)$, the sets S_1 and $(\pi\rho_j)(S_1)$ are mutually disjoint. Since $(\pi \mid Y) \in \Pi'$, $\Pi' = (\pi \mid Y)\Pi' = \{\pi\rho_j \mid Y \colon j = 1, \cdots, k\}$. Let $\pi_j = \pi\rho_j$ $(j = 1, \cdots k)$, so that

$$\Pi' = \{\pi_j \mid Y \colon j = 1, \cdots, k\}. \tag{21}$$

Furthermore, for each j, the sets S_1 and $\pi_j(S_1)$ are mutually disjoint, whence so are $\pi_j^{-1}(S_1)$ and S_1.

If $x \in S$ then, by (19), (20) and the stated properties of π_j, it follows that $\pi_j(x)$ and $\pi_j^{-1}(x)$ lie in $Z - S_1$; thus, again by (20),

$$|u_1(\pi_j(x))| < \epsilon, \quad |u_1(\pi_j^{-1}(x))| < \epsilon \qquad (x \in S; j = 1, \cdots, k). \tag{22}$$

By (21) and (15),

$$\begin{aligned} \sum_{j=1}^{k} \sum_{y \in Y} u_1(y)\overline{u_1(\pi_j(y))} &= \sum_{y \in Y} u_1(y) \sum_{\pi' \in \Pi'} \overline{u_1(\pi'(y))} \\ &= 0, \end{aligned}$$

so there is at least one value of j such that

$$Rl\{\sum_{y\in Y}u_1(y)\overline{u_1(\pi_j(y))}\}\leqq 0.$$

With this value of j,

$$\begin{aligned}\sum_{y\in Y}|u_1(y)-u_1(\pi_j(y))|^2&\\ &=\sum_{y\in Y}\{|u_1(y)|^2+|u_1(\pi_j(y))|^2-2Rl[u_1(y)\overline{u_1(\pi_j(y))}]\}\\ &\geqq\sum_{y\in Y}\{|u_1(y)|^2+|u_1(\pi_j(y))|^2\},\end{aligned}$$

and since $\pi_j(Y)=Y$,

$$\sum_{y\in Y}|u_1(y)-u_1(\pi_j(y))|^2\geqq 2\sum_{y\in Y}|u_1(y)|^2. \tag{23}$$

Since $S\subseteq S_1$, the sets S and $\pi_j^{-1}(S)$ are mutually disjoint, and it is clear from (18) and (19) that neither of them meets Y. From (23), since $R\subseteq Y$, and since u_1 satisfies (13),

$$\begin{aligned}2\sum_{y\in R}|u_1(y)|^2+\sum_{x\in S}|u_1(x)-u_1(\pi_j(x))|^2&\\ +\sum_{x\in S}|u_1(x)-u_1(\pi_j^{-1}(x))|^2&\\ \leqq\sum_{x\in Y\cup S\cup\pi_j^{-1}(S)}|u_1(x)-u_1(\pi_j(x))|^2&\\ \leqq M^2.&\end{aligned}\tag{24}$$

Since $Q=R\cup S$, we may define $a(x)$ and $b(x)$ $(x\in Q)$ by

$$a(x)=u_1(\pi_j(x)),\quad b(x)=u_1(\pi_j^{-1}(x))\qquad (x\in S),$$

and $a(y)=b(y)=0$ $(y\in R)$. It follows from (22) and (24) that (16) and (17) are satisfied. This completes the proof of (14).

It is apparent that u_1 is determined to within an additive constant on each orbit Z. That u_1 is uniquely determined (on each orbit, and hence throughout X) now follows, from the condition $\sum_{x\in Z}|u_1(x)|^2<\infty$ if Z is infinite, and from the fact (deduced from (15)) that u_1 has mean value zero on Z if Z is finite. This completes the proof of Theorem 3.8.

Remark 3.9. If, in Theorem 3.8, the inequalities (13) and (14) are replaced by

$$\sum_{x\in X}|u(x)-u(\pi(x))|^2<\infty\qquad(\pi\in\Pi),$$

and

$$\sum_{x\in X}|u_1(x)|^2<\infty$$

respectively, then the statement so obtained is false. Simple counter-examples occur, for instance, when X is an arbitrary infinite set and Π consists of those permutations of X each of which moves only a finite number of elements.

UNIVERSITY OF PENNSYLVANIA.

REFERENCES.

[1] R. Kadison, "Derivations of operator algsbras," *Annals of Mathematics*, vol. 83 (1966), pp. 280-293.
[2] I. Kaplansky, "Modules over operator algebras," *American Journal of Mathematics*, vol. 75 (1953), pp. 839-859.
[3] P. Miles, "Derivations on B^*-algebras," *Pacific Journal of Mathematics*, vol. 15 (1965), pp. 1359-1366.
[4] F. Murray and J. von Neumann, "On rings of operators, IV," *Annals of Mathematics*, vol. 44 (1943), pp. 716-808.
[5] S. Sakai, "On a conjecture of Kaplansky," *Tôhoku Mathematical Journal*, vol. 12 (1960), pp. 31-33.

THE STRONG SZEGÖ LIMIT THEOREM FOR TOEPLITZ DETERMINANTS.

I. I. HIRSCHMAN, JR.[1]

1. Introduction.

The strong Szegö limit theorem for Toeplitz determinants in one of its more recent forms asserts the following. Let

$$f(\theta)=\sum_{k=-\infty}^{\infty} f^{\wedge}(k)e^{ik\theta}$$

where

$$(1)\qquad \sum_{-\infty}^{\infty}|f^{\wedge}(k)|<\infty,\qquad \sum_{-\infty}^{\infty}|f^{\wedge}(k)|^{2}\,|k|<\infty,$$

and let

$$(2)\qquad f(\theta)\neq 0,\quad\text{and}\quad [\arg f(\theta)]_0^{2\pi}=0.$$

Here $[\arg f(\theta)]_0^{2\pi}$ denotes the change of argument of $f(\theta)$ as θ goes from 0 to 2π. Then if

$$D_n[f]=\det[f^{\wedge}(k-j)]_{k,j=0,\cdots,n},$$

$$G[f]=\exp[\tfrac{1}{2\pi}\int_0^{2\pi}\log f(\theta)\,d\theta],$$

we have

$$(3)\qquad \lim_{n\to\infty}\frac{D_n[f]}{G[f]^{n+1}}=E[f\mid f],$$

where

$$E[f\mid f]=\exp[\tfrac{1}{2}\sum_{-\infty}^{\infty} f_\lambda^{\wedge}(k)f_\lambda^{\wedge}(-k)\,|k|\,].$$

In this formula $f_\lambda^{\wedge}(k)$ is defined by

$$f_\lambda^{\wedge}(k)=\frac{1}{2\pi}\int_0^{2\pi}\log f(\theta)\,e^{-ik\theta}d\theta.$$

Note that the choice of branch for $\log f(\theta)$ does not affect the value of either $G[f]$ or of $E[f\mid f]$.

This striking formula was first obtained by Szegö in [14] under assump-

---

Received May 4, 1965.

[1] Research supported in part by the United States Air Force Office of Scientific Research under Contract No. AF-AFOSR 721-65.

tions considerably stronger than those above. In a series of papers, [1]-[3], Baxter gave a complete and elegant theory of finite section Wiener-Hopf equations in the context of certain Banach algebras of functions on T, the real numbers modulo 2π. As an application he deduced a very general form of the strong Szegö limit theorem. The conditions above, which are somewhat more general than Baxter's, were obtained by the author in [9] and [10]. In these papers Baxter's theory is adapted to apply to a different class of Banach algebras.

Another approach to these questions is due to Kac in [11]. Kac's arguments, as slightly modified by Baxter, see [10], show that (3) holds if (1) is kept unchanged but (2) is replaced by the stronger relation

$$|f^\wedge(0)| > \sum_{k\neq 0} |f^\wedge(k)|. \tag{2'}$$

The condition (2′) is very restrictive. However in the course of his investigations Kac discovered a remarkable combinatorial identity which has led to a whole body of research at the hands of a large number of mathematicians. A particularly accessible treatment of Kac's identity is given in [15].

In the present paper we will give a new and different proof of (3) under assumptions (1) and (2), which makes fundamental use of Kac's argument. The importance of the strong Szegö limit theorem is, I feel, sufficient to justify the publication of a second proof. In addition these new arguments can be applied in some situations where it does not seem possible to use the old ones. In particular they can be used for Toeplitz forms arising from ultraspherical polynomials. For $\nu \geqq 0$ fixed let

$$W_\nu(k,x) = \frac{n!}{(2\nu)_n} C_k{}^\nu(x) \qquad k = 0, 1, \cdots,$$

where $(\alpha)_n = \alpha(\alpha+1)\cdots(\alpha+n-1)$, and where $C_k{}^\nu$ is the standard ultraspherical polynomial of index ν and order k. Let

$$\omega_\nu(k) = \frac{\Gamma(\nu)(k+\nu)\Gamma(k+2\nu)}{\pi^{\frac{1}{2}}\Gamma(\nu+\frac{1}{2})\Gamma(2\nu)k!}.$$

We then have the orthogonality relations

$$\int_1^{-1} W_\nu(k,x) W_\nu(m,x)(1-x^2)^{\nu-\frac{1}{2}} dx = \frac{\delta(k,m)}{\omega_\nu(k)}.$$

We note that for any $\nu \geqq 0$

$$|W_\nu(k,x)| \leqq 1 \qquad k = 0, 1, \cdots;$$

see [4]. Let

$$f(x)=\sum_{k=0}^{\infty} f^{\smile}(k)\, W_\nu(k,x)$$

where

$$\sum_{k=0}^{\infty} |f^{\smile}(k)| < \infty, \qquad \sum_{k=0}^{\infty} |f^{\smile}(k)|^2\, k < \infty. \tag{4}$$

This implies that $f(x)$ is continuous. Suppose furthermore that

$$f(x) \neq 0 \qquad -1 \leqq x \leqq 1. \tag{5}$$

Then if

$$D_n[f] = \det[f^{\wedge}(j,k)]_{j,k=0,\cdots,n}$$

where

$$f^{\wedge}(j,k) = [\omega_\nu(j)\,\omega_\nu(k)]^{\frac{1}{2}} \int_{-1}^{1} f(x)\, W_\nu(j,x)\, W_\nu(k,x)\, (1-x^2)^{\nu-\frac{1}{2}}\, dx,$$

and if

$$G[n,f] = \exp\Big[\int_{-1}^{1} \{\sum_{k=0}^{n} \omega_\nu(k)\, W_\nu(k,x)^2\} \log f(x)\, dx\Big],$$

we will show that

$$\lim_{n\to\infty} \frac{D_n[f]}{G[n,f]} = E[f \mid f]. \tag{6}$$

Here

$$E[f \mid f] = \exp\Big[\tfrac{1}{2}\sum_{1}^{\infty} f_\lambda{}^{\wedge}(k)^2 k\Big]$$

where

$$f_\lambda{}^{\wedge}(k) = \frac{1}{2\pi}\int_0^{2\pi} \log f(\cos\theta)\, e^{-ik\theta}\, d\theta.$$

(Note that since $f(\cos\theta)$ is even $f_\lambda{}^{\wedge}(k) = f_\lambda{}^{\wedge}(-k)$). If we assume in addition that

$$[f(1)-f(x)]/(1-x)^{3/2} \quad \text{and} \quad [f(x)-f(-1)]/(1+x)^{3/2} \tag{7}$$

are integrable then we can rewrite (6) in the simpler form

$$\lim_{n\to\infty} \frac{D_n[f]}{G[f]^{n+1+(2\nu-1)/2}} = [f(1)\cdot f(-1)]^{-(2\nu-1)/4} E[f \mid f] \tag{6'}$$

where

$$G[f] = \exp\Big[\frac{1}{\pi}\int_0^{\pi} \log f(x)\, (1-x^2)^{-\frac{1}{2}}\, dx\Big].$$

In proving (6) an essential step is to show that it is true if (5) is replaced by the stronger assumption

$$(5') \qquad |f^{\vee}(0)| > \sum_{j=1}^{\infty} |f^{\vee}(j)|.$$

This analogue of Kac's result was obtained by Davis and the author in [4]. It depends upon certain Banach algebra properties of ultraspherical polynomials, and thus does not readily extend to more general orthogonal polynomials. The remainder of the proof—in which starting from the fact that (4) and (5′) imply (6) I show (4) and (5) imply (6)—is, on the other hand, fairly general. The passage from (6) to (6′) depends upon the asymptotic behavior of

$$\sum_{k=0}^{n} \omega_\nu(k) W_\nu(k,x)^2$$

as $n \to \infty$, which is studied in [4].

Part I. The Trigonometric Case.

2. Preliminary results. We denote by $\boldsymbol{C}_+$ ($\boldsymbol{C}_-$) the set of those continuous complex valued functions $f(\theta)$, $\theta \in T$, such that $f^{\wedge}(k) = 0$ for $k < 0$ ($k > 0$).

Definition 2a. A continuous complex-valued function $c(\theta)$, $\theta \in T$, is said to belong to $\boldsymbol{R}$ if $c(\theta)$ can be represented in the form

$$(1) \qquad c(\theta) = \frac{d^2}{U(\theta)V(\theta)}$$

where $U(\theta)$, $U(\theta)^{-1} \in \boldsymbol{C}_+$, $V(\theta)$, $V(\theta)^{-1} \in \boldsymbol{C}_-$, and $d^2 = U^{\wedge}(0) = V^{\wedge}(0)$.

Here $U(\theta) = \sum_0^{\infty} U^{\wedge}(k) e^{ik\theta}$, $V(\theta) = \sum_{-\infty}^{0} V^{\wedge}(k) e^{ik\theta}$.

Let $f(\theta) = \sum_{-\infty}^{\infty} f^{\wedge}(k) e^{ik\theta}$ belong to $L^2(T)$. We define

$$E^+(n)f\cdot(\theta) = \sum_{k \geqq n} f^{\wedge}(k) e^{ik\theta},$$

$$E^-(n)f\cdot(\theta) = \sum_{k \leqq n} f^{\wedge}(k) e^{ik\theta},$$

for $n = 0, \pm 1, \pm 2, \cdots$. If

$$\| f \|_2 = \{ \tfrac{1}{2\pi} \int_0^{2\pi} | f(\theta) |^2 \, d\theta \}^{\frac{1}{2}} = \{ \sum_{-\infty}^{\infty} | f^{\wedge}(k) |^2 \}^{\frac{1}{2}}$$

then it is apparent that

$$\| E^+(n) f \|_2 \leqq \| f \|_2, \qquad \| E^-(n) f \|_2 \leqq \| f \|_2.$$

The following results are due to Baxter, see [1]-[3], in a somewhat different setting. In the present setting they are due to Reich [12].

THEOREM 2b. *If* $c \in \mathbf{R}$ *there exists an integer* n_0 *and a constant* A *such that if* $n \geqq n_0$, *then*

$$\| f \|_2 \leqq A \, \| E^+(0)E^-(n)cf \|_2$$

for all $f \in E^+(0)E^-(n)L^2(T)$, *and*

$$\| f \|_2 \leqq A \, \| E^+(-n)E^-(0)cf \|_2$$

for all $f \in E^+(-n)E^-(0)L^2(T)$.

Let $c \in \mathbf{R}$; *we will say that* $c \in \mathbf{R}^*$ *if*

$$\det[c^\wedge(j-k)]_{j,k=0,\cdots,n} \neq 0 \qquad n = 0, 1, \cdots.$$

LEMMA 2c. *Let* $c(\theta) \in \mathbf{R}^*$. *There is an* $\epsilon > 0$ (*depending on* c) *such that if* $b(\theta) \in \mathbf{R}$ *and if* $\| c - b \|_\infty < \epsilon$ *then* $b(\theta) \in \mathbf{R}^*$.

Proof. Let n_0 and A be as in Theorem 2b. Suppose that $\| c - b \|_\infty \cdot A < \frac{1}{2}$; it follows that if $f \in E^+(0)E^-(n)L^2(T)$ and if $n \geqq n_0$ then

$$\begin{aligned} \| f \|_2 &\leqq A \, \| E^+(0)E^-(n)cf \|_2, \\ &\leqq A \, \| E^+(0)E^-(n)bf \|_2 + A \, \| E^+(0)E^-(n)[c-b]f \|_2, \\ &\leqq A \, \| E^+(0)E^-(n)bf \|_2 + A \, \| c-b \|_\infty \, \| f \|_2, \end{aligned}$$

and thus that

$$\| f \|_2 \leqq 2A \, \| E^+(0)E^-(n)bf \|_2.$$

Consequently

$$\det[b^\wedge(j-k)]_{j,k=0,\cdots,n} \neq 0 \qquad \text{for } n \geqq n_0.$$

On the other hand for each $n = 0, 1, \cdots, n_0 - 1$

$$\det[b^\wedge(j-k)]_{j,k=0,\cdots,n} \to \det[c^\wedge(j-k)]_{j,k=0,\cdots,n}$$

as $\| b - c \|_\infty \to 0$, etc.

For $c \in \mathbf{R}^*$ there are for each $n = 0, 1, \cdots$ unique trigonometric polynomials $u(n, \theta) \in E^+(0)E^-(n)L^2(T)$ and $v(n, \theta) \in E^+(-n)E^-(0)L^2(T)$ such that

$$E^+(0)E^-(n)c(\theta)u(n, \theta) = 1,$$

$$E^+(-n)E^-(0)c(\theta)v(n, \theta) = 1.$$

Moreover it is easily seen that

$$d(n^2) = u^\wedge(n,0) = v^\wedge(n,0) = D_{n-1}[c]/D_n[c].$$

Thus $d(n) \neq 0$, $n = 0, 1, \cdots$.

THEOREM 2d. *If $c \in \mathbf{R}^*$ and if $u(n,\theta)$ and $v(n,\theta)$ are defined as above then*

$$\lim_{n\to\infty} \| u(n,\theta) - U(\theta) \|_2 = 0,$$

$$\lim_{n\to\infty} \| v(n,\theta) - V(\theta) \|_2 = 0.$$

Proof. See [12].

It follows in particular that $\lim_{n\to\infty} d(n)^2 = d^2$. We will suppose that $d(n)$ has been so chosen that

$$\lim_{n\to\infty} d(n) = d.$$

We define

$$\psi(n,\theta) = d(n)^{-1} e^{-in\theta} u(n,\theta),$$

$$\phi(n,\theta) = d(n)^{-1} e^{in\theta} v(n,\theta).$$

These are the biorthogonal Szegö polynomials and satisfy the relations

$$(2\pi)^{-1} \int_T \phi(n,\theta) \psi(m,\theta) c(\theta)\, d\theta = \delta(n,m)$$

for $n, m = 0, 1, \cdots$. We write

$$\phi(n,\theta) = \sum_{k=0}^{n} \phi^\wedge(n,k) e^{ik\theta},$$

$$\psi(n,\theta) = \sum_{k=-n}^{0} \psi^\wedge(n,k) e^{ik\theta}.$$

We further define

$$\phi^0(n,z) = \sum_{k=0}^{n} \phi^\wedge(n,k) z^k,$$

$$\psi^0(n,z) = \sum_{k=-n}^{0} \psi^\wedge(n,k) z^k.$$

Similarly we set

$$U^0(z) = \sum_{k=0}^{\infty} U^\wedge(k) z^k \qquad |z| < 1,$$

$$V^0(z) = \sum_{k=-\infty}^{0} V^\wedge(k) z^k \qquad |z| > 1.$$

THEOREM 2e. *If* $c \in \boldsymbol{R}$ *then as* $n \to \infty$

$$\phi^0(n,z) \cong z^n d^{-1} V^0(z) \qquad \text{for } |z| > 1,$$

and

$$\psi^0(n,z) \cong z^{-n} d^{-1} U^0(z) \qquad \text{for } |z| < 1.$$

Proof. This follows immediately from Theorem 2d.

3. A special Szegö theorem. For $f(\theta)$ and $g(\theta)$ such that $\log f(\theta)$ and $\log g(\theta)$ are continuous we formally define

$$G[f] = \exp[(2\pi)^{-1} \int_T \log f(\theta)\, d\theta]$$

and

$$E[f \mid g] = \exp[\tfrac{1}{2} \sum_{k=-\infty}^{\infty} f_\lambda\hat{\ }(k) g_\lambda\hat{\ }(k) \, |k|]$$

where

$$f_\lambda\hat{\ }(k) = (2\pi)^{-1} \int_T \log f(\theta) e^{-ik\theta}\, d\theta,$$

$$g_\lambda\hat{\ }(k) = (2\pi)^{-1} \int_T \log g(\theta) e^{-ik\theta}\, d\theta.$$

In practice f and g will be chosen so that the series for $E[f \mid g]$ is absolutely convergent.

For $c \in \boldsymbol{R}^*$ let $\phi(n,\theta)$ and $\psi(n,\theta)$ be the associated Szegö polynomials. If $f \in L^1(T)$ we set

$$f\tilde{\ }(j,k) = (2\pi)^{-1} \int_T f(\theta) \psi(j,\theta) \phi(k,\theta) c(\theta)\, d\theta, \tag{1}$$

and define

$$D_n[f \mid c] = \det[f\tilde{\ }(j,k)]_{j,k=0,\cdots,n}. \tag{2}$$

Note that if $c(\theta) \equiv 1$ then $\phi(n,\theta) = e^{in\theta}$ and $\psi(n,\theta) = e^{-in\theta}$, and thus

$$D_n[f \mid 1] = \det[f\hat{\ }(j-k)]_{j,k=0,\cdots,n}. \tag{3}$$

The principal result of this section is the following.

THEOREM 3a. *Let* $c \in \boldsymbol{R}^*$ *and let*

$$f^0(z) = \prod_{\alpha=1}^{\nu} (\xi_\alpha - z)$$

where the ξ_α, $\alpha = 1, \cdots, \nu$, *are distinct and lie in* $\{z \mid |z| > 1\}$. *Let* $f(\theta) = f^0(e^{i\theta})$. *Then*

$$\lim_{n\to\infty}\frac{D_n[f \mid c]}{G[f]^{n+1}} = E[f \mid c]^2.$$

Proof. If we define

$$Q_1(n,\theta) = \det\begin{bmatrix} f^\sim(0,0) & f^\sim(1,0) & \cdots & f^\sim(n-1,0) & \phi(0,\theta) \\ f^\sim(0,1) & f^\sim(1,1) & \cdots & f^\sim(n-1,1) & \phi(1,\theta) \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ f^\sim(0,n) & f^\sim(1,n) & \cdots & f^\sim(n-1,n) & \phi(n,\theta) \end{bmatrix},$$

then it is apparent that

$$(4) \qquad (2\pi)^{-1}\int_T Q_1(n,\theta)f(\theta)\psi(k,\theta)c(\theta)\,d\theta = \begin{cases} 0 & 0 \leqq k < n \\ D_n[f \mid c] & k = n. \end{cases}$$

Furthermore,

$$(5) \qquad Q_1^\wedge(n,n) = \phi^\wedge(n,n)D_{n-1}[f \mid c].$$

Let

$$f^0(z)Q_2^0(n,z) = \det\begin{bmatrix} \phi^0(n,z) & \phi^0(n+1,z) & \cdots & \phi^0(n+\nu,z) \\ \phi^0(n,\xi_1) & \phi^0(n+1,\xi_1) & \cdots & \phi^0(n+\nu,\xi_1) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \phi^0(n,\xi_\nu) & \phi^0(n+1,\xi_\nu) & \cdots & \phi^0(n+\nu,\xi_\nu) \end{bmatrix}.$$

It is apparent that $Q_2(n,z)$ is a polynomial in z of degree n and that if $Q_2(n,\theta) = Q_2^0(n,e^{i\theta})$ then

$$(6) \qquad (2\pi)^{-1}\int_T Q_2(n,\theta)f(\theta)\psi(k,\theta)c(\theta)\,d\theta = \begin{cases} 0 & 0 \leqq k < n \\ \Delta_{n+1} & k = n. \end{cases}$$

Here

$$\Delta_n = \det\begin{bmatrix} \phi^0(n,\xi_1) & \cdots & \phi^0(n+\nu-1,\xi_1) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \phi^0(n,\xi_\nu) & \cdots & \phi^0(n+\nu-1,\xi_\nu) \end{bmatrix}.$$

We also have

$$(7) \qquad Q_2^\wedge(n,n) = \Delta_n\phi^\wedge(n+\nu,n+\nu).$$

Comparing (4), (5), (6) and (7) we are led to

$$D_n[f \mid c]\Delta_n \phi^\wedge(n+\nu, n+\nu) = \phi^\wedge(n,n) D_{n-1}[f \mid c]\Delta_{n+1},$$

from which we obtain by multiplication

$$\frac{\Delta_0}{\prod_{k=0}^{\nu-1} \phi^\wedge(k,k)} D_n[f \mid c] = \frac{\Delta_{n+1}}{\prod_{k=1}^{\nu} \phi^\wedge(n+k, n+k)} \tag{8}$$

provided $\Delta_k \neq 0$, $k = 0, 1, \cdots, n$. However the two sides of (8) are polynomials in $\xi_1, \cdots, \xi_\nu$ and since they are equal "in general" (8) holds without exception.

By an evident process of subtracting columns we find that

$$\frac{\Delta_0}{\prod_{k=0}^{\nu-1} \phi^\wedge(k,k)} = \det \begin{bmatrix} 1 & \xi_1 & \cdots & \xi_1^{\nu-1} \\ 1 & \xi_2 & \cdots & \xi_2^{\nu-1} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 1 & \xi_\nu & \cdots & \xi_\nu^{\nu-1} \end{bmatrix} = \prod_{\beta<\alpha} (\xi_\alpha - \xi_\beta). \tag{9}$$

Next it follows from Theorem 2e that as $n \to \infty$

$$\Delta_{n+1} \cong (\xi_1 \cdots \xi_\nu)^{n+1} d^{-\nu} \prod_{\alpha=1}^{\nu} V^0(\xi_\alpha) \det \begin{bmatrix} 1 & \xi_1 & \cdots & \xi_1^{\nu-1} \\ 1 & \xi_2 & \cdots & \xi_2^{\nu-1} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 1 & \xi_\nu & \cdots & \xi_\nu^{\nu-1} \end{bmatrix} \tag{10}$$

$$\cong (\xi_1 \cdots \xi_\nu)^{n+1} d^{-\nu} \prod_{\alpha=1}^{\nu} V^0(\xi_\alpha) \prod_{\beta<\alpha} (\xi_\alpha - \xi_\beta).$$

Finally by the convention introduced after Theorem 2d

$$\lim_{n\to\infty} \phi^\wedge(n+k, n+k) = \lim_{n\to\infty} d(n+k)^{-1} v^\wedge(n+k, 0) = d. \tag{11}$$

Combining (9), (10), and (11) we see that as $n \to \infty$

$$D_n[f \mid c] \cong (\xi_1 \cdots \xi_\nu)^{n+1} d^{-2\nu} \prod_{\alpha=1}^{\nu} V^0(\xi_\alpha). \tag{12}$$

We have

$$\log f(\theta) = \sum_{\alpha=1}^{\nu} \log(\xi_\alpha - e^{i\theta}) = \sum_{\alpha=1}^{\nu} \log \xi_\alpha + \sum_{\alpha=1}^{\nu} \sum_{k=1}^{\infty} -\frac{\xi_\alpha^{-k}}{k} e^{ik\theta}.$$

It follows from this that

$$G[f]=\xi_1\cdots\xi_\nu, \tag{13}$$

and that

$$f_\lambda{}^\wedge(k)=-\sum_{\alpha=1}^{\nu}\frac{\xi_\alpha{}^{-k}}{k}, \qquad k=1,2,\cdots;$$

while $f_\lambda{}^\wedge(k)=0$ if $k<0$. Thus

$$\sum_{k=-\infty}^{\infty}f_\lambda{}^\wedge(k)c_\lambda{}^\wedge(-k)\,|\,k\,|=-\sum_{k=1}^{\infty}\sum_{\alpha=1}^{\nu}\xi_\alpha{}^{-k}c_\lambda{}^\wedge(-k).$$

It follows from (1) of § 2 that with an evident notation

$$c_\lambda{}^\wedge(-k)=-V_\lambda{}^\wedge(-k) \qquad k=1,2,\cdots.$$

Consequently

$$\begin{aligned}\sum_{-\infty}^{\infty}f_\lambda{}^\wedge(k)c_\lambda{}^\wedge(-k)\,|\,k\,| &=\sum_{\alpha=1}^{\nu}\sum_{k=1}^{\infty}\xi_\alpha{}^{-k}V_\lambda{}^\wedge(-k),\\ &=\sum_{\alpha=1}^{\nu}\sum_{k=0}^{\infty}\xi_\alpha{}^{-k}V_\lambda{}^\wedge(k)-2\nu\log d,\\ &=\sum_{\alpha=1}^{\nu}\log V^0(\xi_\alpha)-2\nu\log d,\end{aligned}$$

and thus

$$E[f\mid c]^2=d^{-2\nu}\prod_{\alpha=1}^{\nu}V^0(\xi_\alpha). \tag{14}$$

It is clear that (12), (13), and (14) together prove our theorem

THEOREM 3b. *Let* $c\in \boldsymbol{R}^*$, *and let*

$$f^0(z)=\prod_{\alpha=1}^{\nu}(\xi_\alpha-z^{-1})$$

where the ξ_α, $\alpha=1,\cdots,\nu$, *are distinct and lie in* $\{z\mid |z|>1\}$. *Let* $f(\theta)=f^0(e^{i\theta})$. *Then*

$$\lim_{n\to\infty}\frac{D_n[f\mid c]}{G[f]^{n+1}}=E[f\mid c]^2.$$

Proof. This can be deduced from Theorem 3a by means of the automorphism which sends $h(\theta)$ into $h(-\theta)$, or it can be proved directly by a parallel argument.

4. The multiplication identity.

THEOREM 4a. *If* $c\in \boldsymbol{R}^*$ *and if* $f\in L^1(T)$ *then we have*

$$D_n[fc \mid 1] = D_n[f \mid c] D_n[c \mid 1].$$

Proof. All of the matrices below are $(n+1)\times(n+1)$. It follows from (1) of § 3 that if $h = cf$ then

$$f^{\sim}(j,k) = \sum_{\alpha,\beta=0}^{n} \psi^{\wedge}(j,-\alpha)\, h^{\wedge}(\alpha-\beta)\, \phi^{\wedge}(k,\beta),$$

and thus

$$[f^{\sim}(j,k)] = [\psi^{\wedge}(j,-\alpha)]\,[h^{\wedge}(\alpha-\beta)]\,[\phi^{\wedge}(k,\beta)]^{t}.$$

Taking determinants we find that

$$(1) \qquad D_n[f \mid c] = \det[\psi^{\wedge}(j,-\alpha)]\, D_n[fc \mid 1] \det[\phi^{\wedge}(k,\beta)].$$

For the special choice $f(\theta) \equiv 1$ we find that

$$[\delta(j-k)] = [\psi^{\wedge}(j,-\alpha)]\,[c^{\wedge}(\alpha-\beta)]\,[\phi^{\wedge}(k,\beta)]^{t}$$

and hence

$$(2) \qquad 1 = \det[\psi^{\wedge}(j,-\alpha)]\, D_n[c \mid 1] \det[\phi^{\wedge}(k,\beta)].$$

Inserting (2) into (1) we obtain our desired result.

THEOREM 4b. *Let $c \in \boldsymbol{R}^*$ be such that*

$$\lim_{n\to\infty} \frac{D_n[c \mid 1]}{G[c]^{n+1}} = E[c \mid c].$$

Then if $f(z)$ is as in Theorem 3a (*or Theorem* 3b)

$$\lim_{n\to\infty} \frac{D_n[fc \mid 1]}{G[fc]^{n+1}} = E[fc \mid fc].$$

Proof. Since

$$G[fc] = G[f]G[c]$$

we see using Theorem 4a that

$$\frac{D_n[fc \mid 1]}{G[fc]^{n+1}} = \frac{D_n[f \mid c]}{G[f]^{n+1}} \frac{D_n[c \mid 1]}{G[c]^{n+1}}.$$

Making use of the results of § 3 we see that

$$\lim_{n\to\infty} \frac{D_n[fc \mid 1]}{G[fc]^{n+1}} = E[c \mid c]E[f \mid c]^2 = E[c \mid c]E[f \mid c]^2 E[f \mid f|$$

since $E[f \mid f] = 1$. It is immediately apparent that

$$E[c \mid c]E[f \mid c]^2 E[f \mid f] = E[fc \mid fc]$$

holds identically, and we are done.

5. The strong limit theorem. Let us denote by $\boldsymbol{B}$ the set of all continuous functions $f(\theta)$ on T for which $\| f \| < \infty$ where

$$\| f \| = 3\{ \sum_{k=-\infty}^{\infty} | f^{\wedge}(k) | + [\sum_{k=-\infty}^{\infty} | f^{\wedge}(k) |^2 | k |]^{\frac{1}{2}} \}.$$

It can be verified, see § 10 for a similar demonstration carried out in detail, that $\boldsymbol{B}$ is a Banach algebra.

THEOREM 5a. *If $c(\theta) \in \boldsymbol{B}$, if $c(\theta) \neq 0$ for $\theta \in T$, and if $[\arg c(\theta)]_0^{2\pi} = 0$, then*

$$\lim_{n \to \infty} \frac{D_n[c]}{G[c]^{n+1}} = E[c \mid c].$$

Proof. In view of Kac's theorem quoted in § 1 and Theorem 4b it is enough to prove that we can write $c(\theta)$ in the form

$$c(\theta) = f(\theta) g(\theta) h(\theta)$$

where:

i. $h(\theta) \in \boldsymbol{B}, \qquad | h^{\wedge}(0) | > \sum_{k \neq 0} | h^{\wedge}(k) |, \qquad h \in \boldsymbol{R}^*;$

(1) ii. $g(\theta)$ is as in Theorem 3b and $g(\theta) h(\theta) \in \boldsymbol{R}^*$;

iii. $f(\theta)$ is as in Theorem 3a.

The maximal ideal space of $\boldsymbol{B}$ is T. Using this fact together with the Wiener-Levy theorem we can show, see [9], that

$$c(\theta) = U_1(\theta) V_1(\theta)$$

where

$$U_1(\theta), U_1(\theta)^{-1} \in E^+(0)\boldsymbol{B}, \qquad V_1(\theta), V_1(\theta)^{-1} \in E^-(0)\boldsymbol{B}.$$

Let m be a positive integer and set

$$F(\theta) = E^-(m) U_1(\theta), \qquad G(\theta) = E^+(-m) V_1(\theta).$$

It is evident that if m is large enough then

$$F(\theta)^{-1} \in E^+(0)\boldsymbol{B}, \qquad G(\theta)^{-1} \in E^-(0)\boldsymbol{B}, \tag{2}$$

and thus that $H(\theta) \in \boldsymbol{B}$ where

$$H(\theta) = c(\theta) F(\theta)^{-1} G(\theta)^{-1}.$$

Moreover we see that

$$\lim_{m \to \infty} \| H(\theta) - 1 \| = 0, \qquad \lim_{m \to \infty} \| H(\theta) G(\theta) - U_1(\theta) \| = 0.$$

Let us choose m so large that (2) holds and in addition

$$(3)\qquad |H^{\wedge}(0)| > \sum_{k\neq 0} |H^{\wedge}(k)|, \qquad H \in \boldsymbol{R}^*, \quad HG \in \boldsymbol{R}^*.$$

Here we have used the fact that 1 and $U_1(\theta) \in \boldsymbol{R}^*$, together with Lemma 2c. (Note that $D_n[U_1 \mid 1] \neq 0$ since the corresponding matrix has zero entries above the main diagonal and non-zero entries on the main diagonal.) We can write $F^0(z)$ in the form

$$F^0(z) = r \prod_{\alpha=1}^{\mu} (\zeta_\alpha - z), \qquad r \neq 0,$$

for some integer μ, $0 \leqq \mu \leqq m$. Since $F(\theta)^{-1} \in E^+(0)\boldsymbol{B}$ it follows that $|\zeta_\alpha| > 1$, $\alpha = 1, \cdots, \mu$; however the ζ_α's need not be distinct. Perturb $\{\zeta_\alpha\}_1^\mu$ slightly to obtain $\{\zeta'_\alpha\}_1^\mu$ which are distinct and set

$$f^0(z) = \prod_{\alpha=1}^{\mu} (\zeta'_\alpha - z).$$

Similarly

$$G^0(z) = s \prod_{\beta=1}^{\nu} (\xi_\beta - z^{-1}), \qquad s \neq 0,$$

for some integer ν, $0 \leqq \nu \leqq m$. Since $G(\theta)^{-1} \in E^-(0)\boldsymbol{B}$ it follows that $|\xi_\beta| > 1$, $\beta = 1, \cdots, \nu$. Perturb $\{\xi_\beta\}_1^\nu$ slightly to obtain $\{\xi'_\beta\}_1^\nu$ which are distinct and set

$$g^0(z) = \prod_{\beta=1}^{\nu} (\xi'_\beta - z^{-1}).$$

It is apparent that this can be done in such a way that if

$$h(\theta) = c(\theta) f(\theta)^{-1} g(\theta)^{-1}$$

all of the conditions (1) are satisfied. Note that Lemma 2c is used here again—in order to insure that h and gh are in $\boldsymbol{R}^*$.

Part II. The Ultraspherical Case.

6. The finite section norm inequality. Let $\Omega(dx)$ be a finite positive measure on the Borel sets of $I = \{x \mid |x| \leqq 1\}$.

Definition 6a. Ω belongs to the class $\boldsymbol{S}$ if

$$\int_{-1}^{1} (1 - x^2)^{-\frac{1}{2}} \log \Omega_a(x)\, dx > -\infty$$

where

$$\Omega(dx) = \Omega_s(dx) + \Omega_a(x)\, dx$$

is the decomposition of $\Omega(dx)$ into its singular and absolutely continuous parts.

Let $\{p(k,x)\}_{k=0}^{\infty}$ be the othonormal polynomials corresponding to $\Omega(dx)$. Then $p(k,x) = p_k x^k$ plus terms of lower order. The $p(k,x)$ are uniquely determined by the normalization

$$p_k > 0 \qquad\qquad k = 0, 1, \cdots.$$

In the present situation $\Omega(dx)$ and $\{p(k,x)\}_0^{\infty}$ correspond to $(2\pi)^{-1}\, d\theta$ and $\{e^{ik\theta}\}_{-\infty}^{\infty}$ in the trigonometric case.

For $f(x) \in L^1(\Omega)$ we set

$$f^{\wedge}(k) = \int_{-1}^{1} f(x) p(k,x) \Omega(dx).$$

(Note that the notation in Part II is not the same as in Part I.) If $f(x) \in L^2(\Omega)$ then

$$f(x) = \sum_{k=0}^{\infty} f^{\wedge}(k) p(k,x),$$

the partial sums of the series on the right converging to $f(x)$ in $L^2(\Omega)$. By Parseval's equality

$$\| f \|_2 = \{ \int_{-1}^{1} | f(x) |^2 \Omega(dx) \}^{\frac{1}{2}} = \{ \sum_{k=0}^{\infty} | f^{\wedge}(k) |^2 \}^{\frac{1}{2}}.$$

Let $f \in L^2(\Omega)$. We define the projection operators $\boldsymbol{E}(n)$ by

$$\boldsymbol{E}(n) f \cdot (x) = \sum_{k=0}^{n} f^{\wedge}(k) p(k,x) \qquad\qquad n = 0, 1, \cdots.$$

The principal goal of this section is to obtain an analogue of Theorem 2b. For this purpose we need the following preliminary result

LEMMA 6b. *If* $\Omega \in \boldsymbol{S}$ *and if* $c(x)$, $x \in I$, *is continuous, then*

$$\lim_{k,j \to \infty} \{ \int_{-1}^{1} c(x) p(k,x) p(j,x) \Omega(dx) - \frac{2}{\pi} \int_0^{\pi} \boldsymbol{c}(\theta) \cos k\theta \cos j\theta \, d\theta \} = 0,$$

where $\boldsymbol{c}(\theta) = c(\cos\theta)$.

Proof. Let us write the recurrence formula for the $p(k,x)$ in the form

$$x p(n,x) = A(n) p(n+1,x) + B(n) p(n,x) + C(n) p(n-1,x).$$

Then

$$\lim_{n\to\infty} A(n) = \lim_{n\to\infty} C(n) = \tfrac{1}{2}, \qquad \lim_{n\to\infty} B(n) = 0.$$

This is proved by Szegö [13; § 12.7] in the case $\Omega_s(dx) = 0$. However his demonstration is easily seen to be valid in general. See, for example [7; § 3.4]. At the same time we have

$$\cos\theta \ \cos n\theta = \tfrac{1}{2}\cos(n+1)\theta + \tfrac{1}{2}\cos(n-1)\theta.$$

Our result for $c(x) = x^m$ follows from these remarks together with the orthogonality properties of the various functions involved. The general case can be deduced from this one by an evident approximation argument. See [8] for a different (although closely related) proof which is valid for $c \in L^\infty[-1,1]$.

Definition 6c. $c(x)$ defined on I will be said to belong to $\boldsymbol{R}_I$ if $c(\cos\theta) = \boldsymbol{c}(\theta) \in \boldsymbol{R}$.

Note that if $c \in \boldsymbol{R}_I$ then, since $\boldsymbol{c}(\theta)$ is even, we have

$$\boldsymbol{c}(\theta) = \frac{d^2}{U(\theta)U(-\theta)} \tag{1}$$

where d^2 and $U(\theta)$ satisfy the conditions of Definition 2a.

The analogue of Reich's theorem is the following.

THEOREM 6d. *If $\Omega \in \boldsymbol{S}$ and if $c \in \boldsymbol{R}_I$ there exists an integer n_0 and a constant A such that if $n \geqq n_0$, and if $f \in \boldsymbol{E}(n)L^2(\Omega)$ then*

$$\| f \|_{2,\Omega} \leqq A \ \| \boldsymbol{E}(n) c f \|_{2,\Omega}$$

Since we will, at the same time, be working with functions of θ, $\theta \in T$, and functions of x, $x \in I$, we will (for the time being) write

$$\| f \|_{2,\Omega} = \{ \int_I | f(x) |^2 \Omega(dx) \}^{\frac{1}{2}}, \qquad \| \xi \|_{2,T} = \{ (2\pi)^{-1} \int_T | \xi(\theta) |^2 \, d\theta \}^{\frac{1}{2}}.$$

Proof. As usual let $\boldsymbol{c}(\theta) = c(\cos\theta)$. We assert that there exists an integer n_1 and a constant B such that if $n \geqq n_1$ and if

$$\xi(\theta) \in E^+(-n)E^-(n)L^2(T)$$

then

$$\| \xi(\theta) \|_{2,T} \leqq B \ \| E^+(-n)E^-(n)\boldsymbol{c}(\theta)\xi(\theta) \|_{2,T}.$$

This follows from Theorem 2b and the identity

$$E^+(-n)E^-(n)\boldsymbol{c}(\theta)\xi(\theta) = e^{-in\theta}E^+(0)E^-(2n)\,c(\theta)\xi(\theta)\,e^{in\theta}.$$

By increasing B if necessary we can assume that

$$\| c \|_\infty \leqq B, \quad \| C \|_\infty \leqq B, \tag{2}$$

where $C(x) = 1/c(x)$. Choose polynomials $c_1(x)$ and $C_1(x)$ such that if $c_2(x) = c(x) - c_1(x)$, $C_2(x) = C(x) - C_1(x)$, then

$$\| c_2 \|_\infty \leqq (4B)^{-1}, \qquad \| C_2 \|_\infty \leqq (2[3B^3 + B])^{-1}. \tag{3}$$

Let

$$N = \max[\text{degree } c_1, \text{degree } C_1].$$

If $n \geqq n_1$ and $\xi \in E^+(-n)E^-(n)L^2(T)$ we have

$$\begin{aligned} \| \xi \|_{2,T} &\leqq B \| E^+(-n)E^-(n)\mathbf{c}_1\xi \|_{2,T} + B \| E^+(n)E^-(n)\mathbf{c}_2\xi \|_{2,T} \\ &\leqq B \| E^+(-n)E^-(n)\mathbf{c}_1\xi \|_{2,T} + B \| \mathbf{c}_2 \|_\infty \| \xi \|_{2,T}. \end{aligned}$$

Here $\mathbf{c}_1(\theta) = c_1(\cos\theta)$, etc. Making use of (3) we see that

$$\| \xi \|_{2,T} \leqq \tfrac{4}{3}B \| E^+(-n)E^-(n)\mathbf{c}_1\xi \|_{2,T}.$$

Let us now suppose that

$$\xi(\theta) \in \{E^+(-n)E^-(n) - E^+(-n+N)E^-(n-N)\}L^2(T).$$

Then

$$E^+(-n)E^-(n)\mathbf{c}_1\xi \in \{E^+(-n)E^-(n) - E^+(-n+2N)E^-(n-2N)\}L^2(T),$$

and consequently, if $n \geqq \max(n_1, 2N)$,

$$\begin{aligned} \| \xi \|_{2,T} &\leqq \tfrac{4}{3}B \| \{E^+(-n)E^-(n) - E^+(-n+2N)E^-(n-2N)\}\mathbf{c}_1\xi \|_{2,T} \\ &\leqq \tfrac{4}{3}B \| \{\cdot\ \cdot\ \cdot\}\mathbf{c}\xi \|_{2,T} + \tfrac{4}{3}B \| \{\cdot\ \cdot\ \cdot\}\mathbf{c}_2\xi \|_{2,T}. \end{aligned}$$

Making use of (3) again we find that

$$\| \xi \|_{2,T} \leqq 2B \| \{E^+(-n)E^-(n) - E^+(-n+2N)E^-(n-2N)\}\mathbf{c}\xi \|_{2,T}. \tag{4}$$

If we assume, as we now do, that in addition $\xi(\theta)$ is *even* we may write

$$\xi(\theta) = \sum_{n-N<j\leqq n} \xi^\#(j)\cos j\theta.$$

Furthermore if

$$\eta(\theta) = E^+(-n)\, E^-(n)c(\theta)\xi(\theta)$$

then we may write

$$\eta(\theta) = \sum_{k=0}^{n} \eta^\#(k)\cos k\theta.$$

A simple computation shows that

$$\eta^\#(k) = \sum_{n-N<j\leqq n} m(k,j)\xi^\#(j) \qquad 0 \leqq k \leqq n, \tag{5}$$

where

$$m(k,j)=\frac{2}{\pi}\int_0^{\pi} c(\theta)\cos k\theta\cos j\theta\,d\theta.$$

We have thus proved that given $\xi^{\#}(j)$, $n-N<j\leqq n$, if $\eta^{\#}(k)$ is defined by (5), then

$$\{\sum_{n-N<j\leqq n}|\xi^{\#}(j)|^2\}^{\frac{1}{2}}\leqq 2B\{\sum_{n-2N<k\leqq n}|\eta^{\#}(k)|^2\}^{\frac{1}{2}} \tag{6}$$

provided $n\geqq\max(n_1,2N)$.

Let $f(x)$ be of the form

$$f(x)=\sum_{n-N<j\leqq n} f^{\wedge}(j)p(j,x). \tag{7}$$

Then if

$$g(x)=\boldsymbol{E}(n)c(x)f(x)$$

we may write

$$g(x)=\sum_{k=0}^{n} g^{\wedge}(k)p(k,x).$$

Moreover,

$$g^{\wedge}(k)=\sum_{n-N<j\leqq n} M(k,j)f^{\wedge}(j) \qquad\qquad 0\leqq k\leqq n,$$

where

$$M(k,j)=\int_{-1}^{1} c(x)p(k,x)p(j,x)\Omega(dx).$$

By Lemma 6b

$$\lim_{k,j\to\infty}[M(k,j)-m(k,j)]=0.$$

It follows that there exists an integer $n_0\geqq\max(n_1,2N)$ such that if $f(x)$ satisfies (7) and if $n\geqq n_0$ then

$$\{\sum_{n-N<j\leqq n}|f^{\wedge}(j)|^2\}^{\frac{1}{2}}\leqq 3B\{\sum_{n-2N<j\leqq n}|g^{\wedge}(k)|^2\}^{\frac{1}{2}}.$$

A fortiori if

$$n\geqq n_0,\qquad f\in\{\boldsymbol{E}(n)-\boldsymbol{E}(n-N)\}L^2(\Omega), \tag{8}$$

then

$$\|f\|_{2,\Omega}\leqq 3B\,\|\boldsymbol{E}(n)cf\|_{2,\Omega}. \tag{9}$$

This, of course, is our theorem for a rather restricted f. We will show in what follows that the general case follows from this special one.

Suppose that $n\geqq n_0$ and that $f(x)\in\boldsymbol{E}(n)L^2(\Omega)$. We set

$$f_1(x)=\sum_{j=0}^{n-N} f^\wedge(j)p(j,x), \qquad f_2(x)=\sum_{j=n-N+1}^{n} f^\wedge(j)p(j,x),$$

and we write

$$f(x)c(x)=g(x)+G(x)$$

where $g(x)\in \boldsymbol{E}(n)L^2(\Omega)$ and $G(x)=f(x)c(x)-g(x)$. We have

$$f_1(x)+f_2(x)=C(x)g(x)+C_1(x)G(x)+C_2(x)G(x).$$

Since $\boldsymbol{E}(n-N)C_1G=0$

$$f_1=\boldsymbol{E}(n-N)[Cg+C_2G],$$

$$\begin{aligned}\|f_1\|_{2,\Omega}&\leqq\|C\|_\infty\|g\|_{2,\Omega}+\|C_2\|_\infty\|G\|_{2,\Omega},\\ &\leqq B\|g\|_{2,\Omega}+\|C_2\|_\infty B\|f\|_{2,\Omega}.\end{aligned}\tag{10}$$

Here we have used the trivial inequality $\|G\|_{2,\Omega}\leqq B\|f\|_{2,\Omega}$ which holds because $G=[I-\boldsymbol{E}(n)]cf$. On the other hand

$$cf_2=-cf_1+g+G,$$

$$\boldsymbol{E}(n)cf_2=g-\boldsymbol{E}(n)cf_1,$$

and therefore by (9)

$$\|f_2\|_{2,\Omega}\leqq 3B\|g-\boldsymbol{E}(n)cf_1\|_{2,\Omega}\leqq 3B\|g\|_{2,\Omega}+3B^2\|f_1\|_{2,\Omega},$$

$$\|f\|_{2,\Omega}\leqq\|f_1\|_{2,\Omega}+\|f_2\|_{2,\Omega}\leqq 3B\|g\|_{2,\Omega}+(3B^2+1)\|f_1\|_{2,\Omega}.\tag{11}$$

The insertion of (10) into (11) results in

$$\|f\|_{2,\Omega}\leqq(3B^3+4B)\|g\|_{2,\Omega}+(3B^3+B)\|C_2\|_\infty\|f\|_{2,\Omega}.$$

Making use of the second inequality of (3) we find that

$$\|f\|_{2,\Omega}\leqq A\|g\|_{2,\Omega}$$

if $A=2(3B^3+4B)$.

COROLLARY 6e. *Under the assumptions of Theorem* 6d *given*

$$g\in\boldsymbol{E}(n)L^2(\Omega), \qquad n\geqq n_0,$$

there exists a unique $f\in\boldsymbol{E}(n)L^2(\Omega)$ *such that* $\boldsymbol{E}(n)cf=g$. *Moreover* $\|f\|_{2,\Omega}\leqq A\|g\|_{2,\Omega}$.

Let us set

$$c^\wedge(j,k)=\int_{-1}^{1}c(x)p(j,x)p(k,x)\Omega(dx).$$

If $c \in \boldsymbol{R}_I$ we will say that $c \in \boldsymbol{R}^*_I$ if

$$\det[c^\wedge(j,k)]_{j,k=0,\cdots,n} \neq 0 \qquad n = 0, 1, \cdots.$$

LEMMA 6f. *If $c \in \boldsymbol{R}^*_I$ then there is an $\epsilon > 0$ such that if $b \in \boldsymbol{R}_I$ and if $\| b - c \|_\infty < \epsilon$ then $b \in \boldsymbol{R}^*_I$.*

Proof. See § 2.

7. The asymptotic formulas. We assume throughout that $\Omega \in \boldsymbol{S}$ and $c \in \boldsymbol{R}_I$. By Corollary 6e if $n \geqq n_0$ then there exists a unique polynomial $y(n,x) \in \boldsymbol{E}(n)L^2(\Omega)$ such that

$$\boldsymbol{E}(n)c(x)y(n,x) = p(n,x). \tag{1}$$

LEMMA 7a. *If $y(n,x)$ is defined as above then*

$$\int_{-1}^{1} y(n,x)p(k,x)c(x)\Omega(dx) = 0 \qquad k = 0, \cdots, n-1,$$
$$= 1 \qquad k = n.$$

Proof. By definition $c(x)y(n,x) = p(n,x) + Y(n,x)$ where $Y(n,x)$ is in the linear subspace of $L^2(\Omega)$ spanned by $\{p(m,x)\}_{n+1}^\infty$ and is therefore orthogonal in $L^2(\Omega)$ to $\{p(k,x)\}_0^n$. Thus

$$\int_{-1}^{1} y(n,x)p(k,x)c(x)\Omega(dx)$$
$$= \int_{-1}^{1} p(n,x)p(k,x)\Omega(dx) + \int_{-1}^{1} Y(n,x)p(k,x)\Omega(dx),$$

etc.

Let us write

$$y(n,x) = \sum_{k=0}^{n} y^\wedge(n,k)p(k,x). \tag{2}$$

We set

$$r(n,x) = \sum_{k=0}^{n} U^\wedge(n-k)p(k,x) \tag{3}$$

where $U(\theta) = \sum_0^\infty U^\wedge(k)e^{ik\theta}$ is as in Definition 6c.

The basic result of the present section is the following.

THEOREM 7b. *If $\Omega \in \boldsymbol{S}$ and $c \in \boldsymbol{R}_I$ then*

$$\lim_{n\to\infty} \| y(n,x) - r(n,x) \|_2 = 0.$$

Proof. Let

$$\boldsymbol{E}(n)c(x)r(n,x)=s(n,x)=\sum_{k=0}^{n}s^{\wedge}(n,k)p(k,x).$$

We assert that

$$\lim_{n\to\infty}s^{\wedge}(n,n-k)=\begin{cases}1 & k=0\\ 0 & k=1,2,\cdots.\end{cases} \tag{4}$$

To see this given $\epsilon>0$ choose N so large that

$$\{\sum_{k=N+1}^{\infty}|U^{\wedge}(k)|^{2}\}^{\frac{1}{2}}<\epsilon,$$

and set

$$U'(\theta)=\sum_{k=0}^{N}U^{\wedge}(k)e^{ik\theta},\qquad U''(\theta)=U(\theta)-U'(\theta).$$

Furthermore let

$$r_a(n,x)=\sum_{n-N<k\leqq n}U^{\wedge}(n-k)p(k,x)=\sum_{0\leqq k\leqq N}U^{\wedge}(k)p(n-k,x),$$

$$r_b(n,x)=\sum_{0\leqq k<N-n}U^{\wedge}(n-k)p(k,x),$$

$$s_a(n,x)=\boldsymbol{E}(n)c(x)r_a(n,x),\qquad s_b(n,x)=\boldsymbol{E}(n)c(x)r_b(n,x).$$

With an evident notation we have

$$s_a^{\wedge}(n,n-j)=\sum_{0\leqq k\leqq N}\int_{-1}^{1}U^{\wedge}(k)p(n-k,x)p(n-j,x)c(x)\Omega(dx).$$

Lemma 6b implies that

$$\lim_{n\to\infty}\int_{-1}^{1}p(n-k,x)p(n-j,x)c(x)\Omega(dx)=\boldsymbol{c}^{\wedge}(j-k).$$

Thus

$$\begin{aligned}\lim_{n\to\infty}s_a^{\wedge}(n,n-j)&=\sum_{0\leqq k\leqq N}U^{\wedge}(k)\boldsymbol{c}^{\wedge}(j-k),\\ &=\frac{1}{2\pi}\int_T U'(\theta)\boldsymbol{c}(\theta)e^{-ij\theta}d\theta,\\ &=\frac{1}{2\pi}\int_T U(\theta)\boldsymbol{c}(\theta)e^{-ij\theta}d\theta-\frac{1}{2\pi}\int_T U''(\theta)\boldsymbol{c}(\theta)e^{-ij\theta}\,d\theta.\end{aligned}$$

We have

$$\frac{1}{2\pi}\int_T U(\theta)\boldsymbol{c}(\theta)e^{-ij\theta}d\theta=\begin{cases}1 & j=0\\ 0 & j=1,2,\cdots.\end{cases}$$

An elementary estimate gives

$$\left| \frac{1}{2\pi} \int_T U''(\theta) \boldsymbol{c}(\theta) e^{-ij\theta} d\theta \right| \leqq \| c \|_\infty \| U'' \|_{2,T} \leqq \| c \|_\infty \epsilon.$$

We thus see that if $j \geqq 0$

$$\overline{\lim_{n \to \infty}} | s_a\hat{\ }(n, n-j) - \delta(j) | \leqq \epsilon \| c \|_\infty.$$

On the other hand

$$\| r_b(n, x) \|_{2,\Omega} \leqq \{ \sum_{N+1}^{\infty} | U\hat{\ }(k) |^2 \}^{\frac{1}{2}} < \epsilon,$$

which implies that

$$\| s_b(n, x) \|_{2,\Omega} \leqq \| c \|_\infty \epsilon,$$

and à fortiori

$$| s_b\hat{\ }(n, n-j) | \leqq \| c \|_\infty \epsilon \qquad j = 0, 1, \cdots, n.$$

Combining these results we have shown that

$$\overline{\lim_{n \to \infty}} | s\hat{\ }(n, n-j) - \delta(j) | \leqq 2 \| c \|_\infty \epsilon \qquad j = 0, 1, \cdots.$$

Since ϵ is arbitrary (4) follows.

Let $\eta > 0$ be given. Choose a polynomial $c_1(x)$ such that if $c_2(x) = c(x) - c_1(x)$

$$\| c_2 \|_\infty < \eta, \qquad \| c_1 \|_\infty \leqq \| c \|_\infty,$$

and take M greater than the degree of $c_1(x)$ and so large that

$$\{ \sum_{M<k<\infty} | U\hat{\ }(k) |^2 \}^{\frac{1}{2}} < \eta.$$

Setting

$$r_1(n, x) = \sum_{n-M<k\leqq n} U\hat{\ }(n-k) p(k, x), \quad r_2(n, x) = \sum_{0\leqq k\leqq n-M} U\hat{\ }(n-k) p(k, x),$$

we find that

$$\begin{aligned} \boldsymbol{E}(n-2M) c(x) r(n, x) &= \boldsymbol{E}(n-2M) [c_1 r_1 + c_2 r_1 + c r_2] \\ &= \boldsymbol{E}(n-2M) [c_2 r_1 + c r_2], \end{aligned}$$

and hence, since $\| r_1 \|_{2,\Omega} \leqq \| U \|_{2,T}$, $\| r_2 \|_{2,\Omega} \leqq \eta$, that

$$\| \boldsymbol{E}(n-2M) c(x) r(n, x) \|_{2,\Omega} \leqq \eta \| U \|_{2,T} + \eta \| c \|_\infty \leqq A_1 \eta$$

where A_1 depends only upon c. Since

$$\| \boldsymbol{E}(n-2M) c(x) y(n, x) \|_{2,\Omega} = 0$$

and since η is arbitrary we see that

$$\overline{\lim_{n\to\infty}}\, \| \boldsymbol{E}(n-2M)c(x)\{y(n,x)-r(n,x)\}\|_{2,\Omega} \leqq A_1\eta.$$

On the other hand it follows from (4) that

$$\overline{\lim_{n\to\infty}}\, \| \{\boldsymbol{E}(n)-\boldsymbol{E}(n-2M)\}c(x)\{y(n,x)-r(n,x)\}\|_{2,\Omega} = 0.$$

These two results taken together show that

$$\overline{\lim_{n\to\infty}}\, \| \boldsymbol{E}(n)c(x)\{y(n,x)-r(n,x)\}\|_{2,\Omega} \leqq A_1\eta.$$

Since by Theorem 6d

$$\| y(n,x)-r(n,x)\|_{2,\Omega} \leqq A\ \| \boldsymbol{E}(n)c(x)\{y(n,x)-r(n,x)\}\|_{2,\Omega}$$

if $n \geqq n_0$, we see that

$$\overline{\lim_{n\to\infty}}\, \| y(n,x)-r(n,x)\|_{2,\Omega} \leqq AA_1\eta.$$

Since η is arbitrary our desired result follows.

If $c \in \boldsymbol{R}_I{}^*$ then, since

$$\det[c^\wedge(j,k)]_{j,k=0,\cdots,n} \neq 0 \qquad\qquad n=0,1,\cdots,$$

there is for *each* n, $n=0,1,\cdots$ a unique function

$$y(n,x) = \sum_0^n y^\wedge(n,k)p(k,x)$$

satisfying

$$\int_{-1}^{1} y(n,x)p(k,x)c(x)\Omega(dx) = \begin{cases} 0 & 0 \leqq k < n \\ 1 & k = n. \end{cases}$$

Moreover since

$$y^\wedge(n,n) = \det[c^\wedge(j,k)]_{j,k=0,\cdots,n-1}/\det[c^\wedge(j,k)]_{j,k=0,\cdots,n},$$

if we set $d(n)^2 = y^\wedge(n,n)$, then

$$d(n)^2 \neq 0 \qquad\qquad n=0,1,\cdots.$$

We know from Theorem 7b that

$$\lim_{n\to\infty} d(n)^2 = U^\wedge(0) = d^2.$$

We may clearly assume that $d(n)$ has been so chosen that

$$\lim_{n\to\infty} d(n) = d.$$

If we now set

$$q(n,x)=d(n)^{-1}y(n,x) \qquad n=0,1,\cdots,$$

Then, see Lemma 7a,

$$\int_{-1}^{1} q(k,x)q(j,x)c(x)\Omega(dx)=\delta(k,j) \qquad j,k=0,1,\cdots.$$

Finally we define q_n by

$$q(n,x)=q_nx^n+\text{lower order terms.}$$

The $\{q(n,x)\}_0^\infty$ are the analogue of the $\{\phi(k,x)\}_0^\infty$ and $\{\psi(k,x)\}_0^\infty$ of Part I.

It follows from Szegö [7; § 1.12-§ 1.14] that

$$\Omega_a(\cos\theta)\,|\sin\theta|=|P(\theta)|^2$$

where $P(\theta)\in H^2(T)$ and

$$\log[(2\pi)^{-1}\int_T P(\theta)\,d\theta]=(2\pi)^{-1}\int_T \log P(\theta)\,d\theta.$$

Let

$$P(\theta)=\sum_{k=0}^{\infty}P^\wedge(k)e^{ik\theta} \qquad \text{in } L^2(T).$$

We also set

$$P^0(z)=\sum_{k=0}^{\infty}P^\wedge(k)z^k \qquad |z|<1.$$

Theorem 7c. *If* $\Omega\in\boldsymbol{S}$, *then*

$$p_n\cong(2\pi)^{-\frac{1}{2}}2^n/P^\wedge(0) \qquad n\to\infty,$$

and if $x\notin I$

$$p(n,x)\cong(2\pi)^{-\frac{1}{2}}\xi^n/P^0(\xi^{-1}) \qquad n\to\infty,$$

where $x=\frac{1}{2}(\xi+\xi^{-1})$, $|\xi|>1$. *Such a representation is always possible and is unique.*

Proof. See Szegö [7; § 6.4].

We can now obtain the analogue of Theorem 2e.

Theorem 7d. *Let* $\Omega\in\boldsymbol{S}$ *and let* $c\in\boldsymbol{R}_I{}^*$. *Then*

$$q_n\cong(2\pi)^{-\frac{1}{2}}2^nd/P^\wedge(0) \qquad \textit{as } n\to\infty,$$

and if $x\notin I$

$$q(n,x)\cong(2\pi)^{-\frac{1}{2}}d^{-1}\xi^nU^0(\xi^{-1})/P^0(\xi^{-1}) \qquad \textit{as } n\to\infty.$$

Proof. This follows from Theorem 7b in the form

$$\{\sum_{k=0}^{n} | y^{\wedge}(n,k) - U^{\wedge}(n-k) |^{2}\}^{\frac{1}{2}} \to 0 \qquad \text{as } n \to \infty,$$

and Theorem 7c.

8. The special Szegö theorem. For $f(x)$, $g(x)$ such that $\log f(\theta)$ and $\log g(\theta)$ are continuous we (formally) define

$$G[f] = \exp(\tfrac{1}{\pi} \int_{0}^{\pi} \log f(\theta)\, d\theta),$$

and

$$E[f \mid g] = \exp[\tfrac{1}{2} \sum_{0}^{\infty} f_{\lambda}^{\wedge}(k) g_{\lambda}^{\wedge}(k) k],$$

where

$$f_{\lambda}^{\wedge}(k) = (2\pi)^{-1} \int_{T} \log f(\theta)\, e^{-ik\theta} d\theta, \quad g_{\lambda}^{\wedge}(k) = (2\pi)^{-1} \int_{T} \log g(\theta)\, e^{-ik\theta} d\theta.$$

For $c \in \boldsymbol{R}_{I}^{*}$ let $\{q(n,x)\}_{0}^{\infty}$ be the generalized orthogonal polynomials with respect to $c(x)\Omega(dx)$, as defined in § 7. We set

$$f^{\sim}(j,k) = \int_{-1}^{1} f(x) q(j,x) q(k,x) c(x) \Omega(dx),$$

and define

$$D[f \mid c] = \det[f^{\sim}(j,k)]_{j,k=0,\cdots,n}.$$

If $c = 1$ then $q(n,x) = p(n,x)$ and

$$D_{n}[f \mid 1] = \det[f^{\wedge}(j,k)]_{j,k=0,\cdots,n}.$$

THEOREM 8a. *Let* $c \in \boldsymbol{R}_{I}^{*}$ *ond let*

$$f(x) = \prod_{\alpha=1}^{\mu} (x_{\alpha} - x)$$

where the x_{α}*'s are distinct and lie in the complement of* I. *Then*

$$\lim_{n \to \infty} \frac{D_{n}[f \mid c]}{G[f]^{n+1}} = E[f \mid c]^{2} E[f \mid f] E[f \mid \Omega_{a}]^{2}.$$

Proof. If we define

$$Q_{1}(n,x) = \det \begin{bmatrix} f^{\sim}(0,0) & f^{\sim}(1,0) & \cdots & f^{\sim}(n-1,0) & q(0,x) \\ f^{\sim}(0,1) & f^{\sim}(1,1) & \cdots & f^{\sim}(n-1,1) & q(1,x) \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ f^{\sim}(0,n) & f^{\sim}(1,n) & \cdots & f^{\sim}(n-1,n) & q(n,x) \end{bmatrix}$$

then $Q_1(n,x)$ is a polynomial of order n,

$$(1) \qquad \int_{-1}^{1} Q_1(n,x)q(k,x)f(x)c(x)\Omega(dx) = \begin{cases} 0 & 0 \leqq k < n \\ D_n[f \mid c] & k = n, \end{cases}$$

and, if $Q_1(n,x) = Q_{1,n}x^n +$ terms of lower order,

$$(2) \qquad Q_{1,n} = q_n D_{n-1}[f \mid c].$$

We also set

$$f(x)Q_2(n,x) = \det \begin{bmatrix} q(n,x) & q(n+1,x) & \cdots & q(n+\mu,x) \\ q(n,x_1) & q(n+1,x_1) & \cdots & q(n+\mu,x_1) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ q(n,x_\mu) & q(n+1,x_\mu) & \cdots & q(n+\mu,x_\mu) \end{bmatrix}.$$

Clearly $Q_2(n,x)$ is a polynomial of degree n and

$$(1') \qquad \int_{-1}^{1} Q_2(n,x)q(k,x)f(x)c(x)\Omega(dx) = \begin{cases} 0 & 0 \leqq k < n \\ \Delta_{n+1} & k = n, \end{cases}$$

and if $Q_2(n,x) = Q_{2,n}x^n +$ terms of lower order

$$(2') \qquad Q_{2,n} = \Delta_n q_{n+\mu}.$$

Here

$$\Delta_n = \det \begin{bmatrix} q(n,x_1) & \cdots & q(n+\mu-1,x_1) \\ q(n,x_2) & \cdots & q(n+\mu-1,x_2) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ q(n,x_\mu) & \cdots & q(n+\mu-1,x_\mu) \end{bmatrix}.$$

Comparing (1), (2), (1′), and (2′) we see that

$$D_n[f \mid c]\Delta_n q_{n+\mu} = q_n D_{n-1}[f \mid c]\Delta_{n+1},$$

from which it follows that

$$(3) \qquad [\prod_{k=0}^{\mu-1} q_k]^{-1} \Delta_0 D_n[f \mid c] = [\prod_{k=1}^{\mu} q_{n+k}]^{-1} \Delta_{n+1}$$

provided no Δ_k, $k = 0, \cdots, n$ is zero. Since both sides of (3) are polynomials in $x_1, \cdots, x_\mu$ an easy argument shows that (3) holds without exception. As in § 2 we see that

$$(4)\qquad \begin{aligned}[\prod_{k=0}^{\mu-1} q_k]^{-1}\Delta_0 &= \prod_{\beta<\alpha}(x_\alpha - x_\beta),\\ &= 2^{-\mu(\mu-1)/2}\prod_{\beta<\alpha}(\xi_\alpha-\xi_\beta)(1-\xi_\alpha^{-1}\xi_\beta^{-1}),\end{aligned}$$

where $x_\alpha = \frac{1}{2}(\xi_\alpha+\xi_\alpha^{-1})$, $|\xi_\alpha| > 1$, $\alpha = 1,\cdots,\mu$. Next it follows that Theorem 7d that

$$(5)\qquad \Delta_{n+1} \cong (2\pi)^{-\mu/2}d^{-\mu}(\xi_1\cdots\xi_\mu)^{n+1}\prod_{\alpha=1}^{\mu}[U^0(\xi_\alpha^{-1})/P^0(\xi_\alpha^{-1})]\prod_{\beta<\alpha}(\xi_\alpha-\xi_\beta)$$

$$\text{as } n\to\infty.$$

A second application of Theorem 7d gives

$$(6)\qquad [\prod_{k=1}^{\mu} q_{n+k}] \cong (2\pi)^{-\mu/2}2^{n\mu+\mu(\mu+1)/2}[dP^\wedge(0)^{-1}]^{\mu}.$$

Collecting these estimates we find that

$$(7)\qquad D_n[f \mid c] \cong \frac{(\xi_1\cdots\xi_\mu)^{n+1}P^\wedge(0)^\mu\prod_{\alpha=1}^{\mu}U^0(\xi_\alpha^{-1})}{2^{(n+1)\mu}d^{2\mu}\prod_{\alpha=1}^{\mu}P^0(\xi_\alpha^{-1})\prod_{\alpha<\beta}(1-\xi_\alpha^{-1}\xi_\beta^{-1})}$$

We have

$$\begin{aligned}\log \boldsymbol{f}(\theta) &= \sum_{\alpha=1}^{\mu}\log(x_\alpha-\cos\theta)\\ &= \sum_{\alpha=1}^{\mu}\{\log(\xi_\alpha/2)+\log(1-e^{i\theta}\xi_\alpha^{-1})+\log(1-e^{-i\theta}\xi_\alpha^{-1})\}\end{aligned}$$

It follows that

$$G[f] = \exp\{\sum_{\alpha=1}^{\mu}\log(\xi_\alpha/2)\} = \{\xi_1\cdots\xi_\mu\}2^{-\mu},$$

and that

$$\boldsymbol{f}_\lambda^\wedge(k) = -\frac{1}{k}\sum_{\alpha=1}^{\mu}\xi_\alpha^{-k}\qquad\qquad k=1,2,\cdots.$$

We have

$$\begin{aligned}\log E[f \mid c] &= \tfrac{1}{2}\sum_{k=1}^{\infty}k\boldsymbol{f}_\lambda^\wedge(k)\mathbf{c}_\lambda^\wedge(k)\\ &= -\tfrac{1}{2}\sum_{\alpha=1}^{\mu}\sum_{k=1}^{\infty}\xi_\alpha^{-k}\mathbf{c}_\lambda^\wedge(k)\\ &= \tfrac{1}{2}\sum_{\alpha=1}^{\mu}\sum_{k=1}^{\infty}\xi_\alpha^{-k}\boldsymbol{U}_\lambda^\wedge(k)\end{aligned}$$

$$= \tfrac{1}{2} \sum_{\alpha=1}^{\mu} \sum_{k=0}^{\infty} \xi_\alpha^{-k} \boldsymbol{U}_\lambda\hat{\ } (k) - \mu \log d.$$

Consequently

$$E[f \mid c]^2 = \prod_{\alpha=1}^{\mu} U^0(\xi_\alpha^{-1})/d^{2\mu}.$$

Similarly

$$\log E[f \mid f] = \tfrac{1}{2} \sum_{\alpha,\beta=1}^{\mu} \sum_{k=1}^{\infty} k^{-1} \xi_\alpha^{-k} \xi_\beta^{-k},$$

$$= \tfrac{1}{2} \sum_{\alpha,\beta=1}^{\mu} -\log(1 - \xi_\alpha^{-1}\xi_\beta^{-1}),$$

so that

$$E[f \mid f] = [\prod_{\alpha<\beta} (1 - \xi_\alpha^{-1}\xi_\beta^{-1})]^{-1} [\prod_{\alpha=1}^{\mu} (1 - \xi_\alpha^{-2})]^{-\frac{1}{2}}.$$

A similar computation shows that

$$E[f \mid \Omega_\alpha]^2 = P\hat{\ }(0)^{\mu} \prod_{\alpha=1}^{\mu} (1 - \xi_\alpha^{-2})^{\frac{1}{2}} / \prod_{\alpha=1}^{\mu} P^0(\xi_\alpha^{-1}).$$

Note that

$$\log \Omega_\alpha(\cos\theta) = \log P(\theta) + \log P(-\theta) - \log |\sin\theta|$$

and

$$-\log|\sin\theta| = \log 2 + \sum_{\substack{k=-\infty \\ k\neq 0}}^{k=\infty} e^{2ik\theta} |2k|^{-1}.$$

Combining these results we obtain our desired conclusion.

Corollary. 8b. *If $f(x)$ is as in Theorem* 8a *then*

$$\lim_{n\to\infty} \frac{D_n[f \mid 1]}{G[f]^{n+1}} = E[f \mid f] E[f \mid \Omega_a]^2.$$

Proof. Take $c = 1$ in Theorem 8a.

This result is the strong Szegö limit theorem relative to the $\{p(k,x)\}_0^\infty$, but for a very special class of functions f.

9. The multiplication identity. The demonstrations of the following results parallel closely those given in § 4 and are omitted.

Theorem 9a. *If $f \in \boldsymbol{R}_I^*$ and if $g \in L^1(\Omega)$ then we have*

$$D_n[fg \mid 1] = D_n[g \mid f] D_n[f \mid 1].$$

Theorem 9b. *Let $c \in \boldsymbol{R}_I^*$ be such that*

$$\lim_{n\to\infty} \frac{D_n[c \mid 1]}{G[c]^{n+1}} = E[c \mid c] E[c \mid \Omega_a]^2.$$

Then if $f(x)$ *is as in Theorem* 8a *we have*

$$\lim_{n\to\infty}\frac{D_n[fc\mid 1]}{G[fc]^{n+1}}=E[fc\mid fc]E[fc\mid \Omega_a]^2.$$

10. The strong limit theorem. Fix ν and let $\{W_\nu(n,x)\}_0^\infty$ be the ultraspherical polynomials nomalized as in §1. We define $\boldsymbol{B}_\nu$ to be the collection of those continuous functions $f(x)$, $x\in I$, of the form

$$f(x)=\sum_{k=0}^{\infty}f^{\smile}(k)\,W_\nu(k,x),$$

for which $\|f\|$ is finite where

$$\|f\|=a(\nu)\{\sum_{1}^{\infty}|f^{\smile}(k)|+[\sum_{0}^{\infty}|f^{\smile}(k)|^2k]^{\frac{1}{2}}\}.$$

Here $a(\nu)$ is a certain positive constant defined below.

LEMMA 10a. *$\boldsymbol{B}_\nu$ is a Banach algebra.*

Proof. It is obvious that $\boldsymbol{B}_\nu$ is a Banach space. We recall that

$$W_\nu(k,x)\,W_\nu(j,x)=\sum_{n=0}^{\infty}c_\nu(k,j,n)\,W_\nu(n,x)\,\omega_\nu(n)$$

where

$$(1)\qquad c_\nu(k,j,n)\geqq 0,\qquad \sum_{n=0}^{\infty}c_\nu(k,j,n)\,\omega_\nu(n)=1,$$
$$c_\nu(k,j,n)=0\qquad \text{if } n<|k-j| \text{ or } n>k+j.$$

The familiar explicit formula for $c_\nu(k,j,n)$ is given in §2 of [4]. Let

$$\boldsymbol{N}_1(f)=\sum_{0}^{\infty}|f^{\smile}(k)|,\qquad \boldsymbol{N}_2(f)=\{\sum_{0}^{\infty}|f^{\smile}(k)|^2k\}^{\frac{1}{2}}.$$

If $f,g\in\boldsymbol{B}_\nu$ and if $h=fg$ then it follows from the above that

$$h^{\smile}(n)=\sum_{j,k=0}^{\infty}f^{\smile}(j)\,g^{\smile}(k)\,c_\nu(k,j,n)\,\omega_\nu(n),$$

$$|h^{\smile}(n)|\leqq\sum_{j,k=0}^{\infty}|f^{\smile}(j)|\,|g^{\smile}(k)|\,c_\nu(k,j,n)\,\omega_\nu(n).$$

Using (1) we see that

$$\boldsymbol{N}_1(h)\leqq\boldsymbol{N}_1(f)\boldsymbol{N}_1(g).$$

Since $c_\nu(k,j,n)=0$ if $n>k+j$ we have

$$| h^{\vee}(n) | \leqq \sum_{k \geqq n/2}^{\infty} \sum_{j=0}^{\infty} | f^{\vee}(j) g^{\vee}(k) | \, c_\nu(k,j,n) \omega_\nu(n)$$

$$+ \sum_{j \geqq n/2}^{\infty} \sum_{k=0}^{\infty} | f^{\vee}(j) g^{\vee}(k) | \, c_\nu(k,j,n) \omega_\nu(n),$$

and hence

$$| h^{\vee}(n) |^2 \leqq 2I_1 + 2I_2,$$

where

$$I_1(n) = \{ \sum_{k \geqq n/2} \sum_{j=0}^{\infty} | f^{\vee}(j) g^{\vee}(k) | \, c_\nu(k,j,n) \omega_\nu(n) \}^2,$$

$$I_2(n) = \{ \sum_{j \geqq n/2} \sum_{k=0}^{\infty} | f^{\vee}(j) g^{\vee}(k) | \, c_\nu(k,j,n) \omega_\nu(n) \}^2.$$

By Schwarz's inequality

$$I_1(n) \leqq \{ \sum_{j,k=0}^{\infty} | f^{\vee}(j) | \, c_\nu(k,j,n) \omega_\nu(k) \}$$

$$\cdot \{ \sum_{k \geqq n/2} \sum_{j=0}^{\infty} | f^{\vee}(j) | \, | g^{\vee}(k) |^2 \, c_\nu(k,j,n) \omega_\nu(n)^2 \omega_\nu(k)^{-1} \},$$

$$nI_1(n) \leqq N_1(f) \sum_{k \geqq n/2} \sum_{j=0}^{\infty} | f^{\vee}(j) | \, | g^{\vee}(k) |^2 \, c_\nu(k,j,n) n \omega_\nu(n)^2 \omega_\nu(k)^{-1}.$$

Since $\omega_\nu(n) \cong \xi_\nu n^{2\nu}$ as $n \to \infty$ where ξ_ν is a positive constant depending only on ν, it follows that there is a constant $A(\nu) \geqq 1$ such that

$$n \omega_\nu(n) \omega_\nu(k)^{-1} \leqq A(\nu) k \qquad \text{if } k \geqq n/2.$$

Thus

$$nI_1(n) \leqq A(\nu) N_1(f) \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} | f^{\vee}(j) | \, k \, |(g^{\vee}(k) |^2 \, c_\nu(k,j,n) \omega_\nu(n)$$

and hence using (1)

$$\sum_{0}^{\infty} nI_1(n) \leqq A(\nu) N_1(f)^2 N_2(g)^2.$$

A similar argument shows that

$$\sum_{n=0}^{\infty} nI_2(n) \leqq A(\nu) N_1(g)^2 N_2(f)^2.$$

Combining these results we see that

$$N_2(h)^2 \leqq 2A(\nu) [N_1(f)^2 N_2(g)^2 + N_2(f)^2 N_1(g)^2],$$

$$N_2(h) \leqq [2A(\nu)]^{\frac{1}{2}} [N_1(f) N_2(g) + N_2(f) N_1(g)],$$

$$\boldsymbol{N}_1(h)+\boldsymbol{N}_2(h)\leqq[2A(\nu)]^{\frac{1}{2}}[\boldsymbol{N}_1(f)\boldsymbol{N}_2(g)+\boldsymbol{N}_2(f)\boldsymbol{N}_1(g)+\boldsymbol{N}_1(f)\boldsymbol{N}_1(g)],$$

$$\leqq[2A(\nu)]^{\frac{1}{2}}[\boldsymbol{N}_1(f)+\boldsymbol{N}_2(f)][\boldsymbol{N}_1(g)+\boldsymbol{N}_2(g)].$$

Consequently if $a(\nu)=[2A(\nu)]^{\frac{1}{2}}$ and if

$$\| f \| = a(\nu)[\boldsymbol{N}_1(f)+\boldsymbol{N}_2(f)]$$

we have

$$\| fg \| \leqq \| f \| \, \| g \|$$

and Lemma 10a is proved.[2]

For $\nu \geqq 0$ fixed we set

$$\Omega(dx)=(1-x^2)^{\nu-\frac{1}{2}}\,dx,$$

$$p(k,x)=W_\nu(k,x)\,\omega_\nu(k)^{\frac{1}{2}}.$$

Then

$$D_n[f \mid 1]=\det[f^{\wedge}(j,k)]_{j,k=0,\cdots,n}$$

where

$$f^{\wedge}(j,k)=[\omega_\nu(j)\,\omega_\nu(k)]^{\frac{1}{2}}\int_{-1}^{1} f(x)\,W_\nu(j,x)\,W_\nu(k,x)\,(1-x^2)^{\nu-\frac{1}{2}}\,dx.$$

As in § 1 we set

$$G[n,f]=\exp\Bigl[\int_{-1}^{1}\Bigl\{\sum_{k=0}^{n}\omega_\nu(k)\,W_\nu(k,x)^2\Bigr\}\log f(x)\,(1-x^2)^{\nu-\frac{1}{2}}\,dx\Bigr].$$

Compare in this connection § 5 of [4].

Theorem 10b. *If* $f(x)\in\boldsymbol{B}_\nu$ *and if* $|f^{\vee}(0)|>\sum_{k=1}^{\infty}|f^{\vee}(k)|$ *then*

$$\lim_{n\to\infty}\frac{D_n[f \mid 1]}{G[n,f]}=E[f \mid f].$$

Proof. This is demonstrated in [4; § 5].

We can now prove the principal result of Part II.

Theorem 10c. *If* $c\in\boldsymbol{B}_\nu$ *and if* $c(x)\neq 0$ *for* $x\in I$ *then*

$$\lim_{n\to\infty}\frac{D_n[c \mid 1]}{G[n,c]}=E[c \mid c]. \tag{2}$$

Proof. Just as in the proof of Theorem 5a we can show that

---

[2] The norm $\|\cdot\|$ does not have the property that $\| 1 \| = 1$. However a familiar argument shows that $\|\cdot\|$ can be replaced by and equivalent norm which does.

$$c(x) = f(x)g(x)$$

where:

i. $f(x) \in \boldsymbol{B}_\nu,\ |f^\vee(0)| > \sum_{k=1}^{\infty} |f^\vee(k)|, f \in \boldsymbol{R}_I^*$;

ii. $g(x) = \prod_{\alpha=1}^{\mu} (x_\alpha - x)$ where the x_α's are distinct and lie in the complement of I.

By Theorem 9a.

$$D_n[c] = D_n[g \mid f] D_n[f \mid 1].$$

Since

$$G[n, c] = G[n, f]\, G[n, g]$$

it follows that

$$\frac{D_n[c]}{G[n, c]} = \frac{D_n[f \mid 1]}{G[n, f]} \cdot \frac{D_n[g \mid f]}{G[g]^{n+1}} \cdot \frac{G[g]^{n+1}}{G[n, g]}.$$

Here

$$G[g] = \exp\{\tfrac{1}{\pi} \int_{-1}^{1} (1 - x^2)^{-\frac{1}{2}} \log g(x)\, dx\}.$$

By Theorem 10b

$$\lim_{n\to\infty} \frac{D_n[f \mid 1]}{G[n, f]^{n+1}} = E[f \mid f].$$

Theorem 8a implies that

$$\lim_{n\to\infty} \frac{D_n[g \mid f]}{G[g]^{n+1}} = E[g \mid f]^2 E[g \mid g] E[g \mid \Omega_a]^2$$

where $\Omega_a(x) = (1 - x^2)^{\nu - \frac{1}{2}}$. Since $g(x)$ is a polynomial $\log g(x)$ trivially satisfies the conditions (7) of § 1, and therefore by Theorem 3c of [4]

$$\lim_{n\to\infty} \frac{G[g]^{n+1}}{G[n; g]} = \frac{[g(1)g(-1)]^{(2\nu-1)/4}}{G[g]^{(2\nu-1)/2}}.$$

We have

$$\begin{aligned} \Omega_a(\cos\theta) &= |\sin\theta|^{2\nu-1}, \\ \log \Omega_a(\cos\theta) &= (2\nu - 1)\log|\sin\theta| \\ &= -(2\nu - 1)\{\log 2 + \sum_{k\neq 0} e^{2ik\theta} |2k|^{-1}\} \end{aligned}$$

It follows that

$$E[g \mid \Omega_a]^2 = [\exp\{\sum_{k=1}^{\infty} g_\lambda{}^\wedge(2k)\}]^{-\nu+\frac{1}{2}} = [g(1)g(-1)]^{-(2\nu-1)/4} G[g]^{(2\nu-1)/2}.$$

Combining these results we obtain

$$\lim_{n\to\infty}\frac{D_n[c]}{G[n,c]}E[f\mid f]E[f\mid g]^2\ \ E[g\mid g]=E[c\mid c].$$

as desired.

Corollary 10d. *If* $c\in\boldsymbol{B}_\nu$, *if* $c(x)\neq 0$ *for* $x\in I$, *and if*

$$[c(1)-c(x)]/(1-x)^{3/2}\ \textit{and}\ [c(x)-c(-1)]/(1+x)^{3/2}$$

are integrable then

$$\lim_{n\to\infty}\frac{D_n[c]}{G[c]^{n+(2\nu-1)/2}}=[c(1)c(-1)]^{-(2\nu-1)/4}E[c\mid c].$$

Proof. Note that our assumptions imply that $[\log c(1)-\log c(x)]/(1-x)^{3/2}$ and $[\log c(x)-\log c(-1)]/(1+x)^{3/2}$ are integrable.

Part III. Miscellaneous Results.

11. An asymptotic formula. Let $\Omega\in\boldsymbol{S}$ and let $\{p(k,x)\}_0^\infty$ be the corresponding orthonormal polynomials normalized as in § 6. In this section we will apply the ideas of Part II in order to study the asymptotic behavior of $\sum_0^n p(k,x)^2$ as $n\to\infty$.

We define

$$c(j,k,l)=\int_{-1}^{1}p(j,x)p(k,x)p(l,x)\Omega(dx),$$

and also

$$p(j,\cos\theta)=\sum_{k=-j}^{j}E(k,j)e^{ik\theta}.$$

Lemma 11a. *Under the above assumptions*

$$\lim_{n\to\infty}c(a+N,j,b+N)=E(b-a,j).$$

Proof. This is a special case of Lemma 6b.

Let

$$\log\Omega_\alpha(\cos\theta)=\sum_{-\infty}^{\infty}\Omega_\lambda{}^\wedge(k)e^{ik\theta}.$$

Note that $\Omega_\lambda{}^\wedge(k)=\Omega_\lambda{}^\wedge(-k)$.

Theorem 11b. *Let* $\Omega\in\boldsymbol{S}$ *and let* $\{p(k,x)\}_0^\infty$ *be the corresponding orthonormal polynomials. Then for any polynomial* $f(x)$

$$\int_{-1}^{1} \{\sum_{k=0}^{n} p(k,x)^2\} f(x)\Omega(dx)$$
$$= \frac{n+1}{\pi} \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} f(x)\,dx + \sum_{k=1}^{\infty} f^{\wedge}(k)\Omega_{\lambda}{}^{\wedge}(k)k + o(1)$$

as $n \to \infty$. *Here* $f(\cos\theta) = \sum_{-\infty}^{\infty} f^{\wedge}(k) e^{ik\theta}$.

Proof. Let

$$f(x) = \sum_{j=0}^{s} f^{\wedge}(j)p(j,x), \qquad f^{\wedge}(j) = \int_{-1}^{1} f(x)p(j,x)\Omega(dx).$$

Suppose that $\rho > 0$ is so small that if $\| f \|_{\infty} = \sup | f(x) |$ for $x \in I$ then

$$\rho \| f \|_{\infty} < 1,$$

and that in addition if $| z | = \rho$, $z = z_1 + iz_2$, then $F(z,x) = 1 - zf(x)$ has, as a polynomial in x, no multiple roots. In order to make our formula shorter we will sometimes write F for $F(z,x)$. As a special case of Theorem 8a we have

$$(1) \quad \lim_{n\to\infty} \{\log D_n[F] - (n+1)\log G[F]\} = \log E[F \mid F] + 2\log E[F \mid \Omega_\alpha]$$

for $| z | = \rho$, and it is easy to see that this holds uniformly for $| z | = \rho$. Since all the functions involved are analytic for $| z | \leqq \rho$, (1) holds uniformly for $| z | \leqq \rho$.

Let $\{\lambda_\nu(n)\}_{\nu=0}{}^{n}$ be the eigenvalues of $A_n(f) = [f^{\wedge}(j,k)]_{j,k=0,\cdots,n}$. Then for any non-negative integer α

$$\sum_{\nu=0}^{n} \lambda_n(n)^{\alpha} = \operatorname{tr}[A_n(f)^{\alpha}].$$

We have

$$\operatorname{tr}[A_n(f)^{\alpha}] = \sum_{k_1,\cdots,k_\alpha=0}^{n} f^{\wedge}(k_1,k_2) f^{\wedge}(k_2,k_3), \cdots, f^{\wedge}(k_\alpha,k_1),$$

and since

$$f^{\wedge}(k_r,k_{r+1}) = \sum_{j_r=0}^{s} c(k_r,j_r,k_{r+1}) f^{\wedge}(j_r),$$

it follows that

$$\operatorname{tr}[A_n(f)^{\alpha}] = \sum_{j_1,\cdots,j_\alpha=0}^{s} f^{\wedge}(j_1)\cdots f^{\wedge}(j_\alpha)$$
$$\times \sum_{k_1,\cdots,k_\alpha=0}^{n} c(k_1,j_1,k_2)c(k_2,j_2,k_3)\cdots c(k_\alpha,j_\alpha,k_1).$$

On the other hand

$$\int_{-1}^{1} f(x)^\alpha \{ \sum_{k=0}^{n} p(k,x)^2 \} \Omega(dx)$$

$$= \sum_{j_1,\cdots,j_\alpha=0}^{s} f^\wedge(j_1) \cdots f^\wedge(j_\alpha)$$

$$\times \sum_{k=0}^{n} \int_{-1}^{1} p(k,x) p(j_1,x) \cdots p(j_\alpha,x) p(k,x) \Omega(dx)$$

$$= \sum_{j_1,\cdots,j_\alpha=0}^{s} f^\wedge(j_1) \cdots f^\wedge(j_\alpha)$$

$$\times \sum_{k_1=0}^{n} \sum_{k_2,\cdots,k_s=0}^{\infty} c(k_1,j_1,k_2) c(k_2,j_2,k_3) \cdots c(k_\alpha,j_\alpha,k_1).$$

For the easy justification of this last step see [4; § 2]. We now have

$$\mathrm{tr}[A_n(f)^\alpha] - \int_{-1}^{1} f(x)^\alpha \{ \sum_{0}^{n} p(k,x)^2 \} \Omega(dx)$$

$$= - \sum_{j_1,\cdots,j_\alpha=0}^{s} f^\wedge(j_1) \cdots f^\wedge(j_\alpha) \sum_{Q(n)} c(k_1,j_1,k_2) \cdots c(k_\alpha,j_\alpha,k_1)$$

where $Q(n)$ consists of all α-tuples $(k_1,\cdots,k_\alpha)$ for which $0 \leqq k_1 \leqq n$, $0 \leqq k_\beta < \infty$, $\beta = 2,\cdots,\alpha$ and $k_\beta > n$ for at least one $\beta = 2,\cdots,\alpha$. Let $h_1 = k_2 - k_1$, $h_2 = k_3 - k_2, \cdots, h_\alpha = k_\alpha - k_1$. Also let $k = k_1 - n$. We have

$$\sum_{Q(n)} c(k_1,j_1,k_2) \cdots c(k_\alpha,j_\alpha,k_1)$$

$$= \sum c(k+n, j_1, k+n+h_1) c(k+n+h_1, j_1, k+n+h_1+h_2) \cdots$$

$$\times c(k+n+h_1+\cdots+h_{\alpha-1}, j_\alpha, k+n)$$

where the summation on the right is extended over those indices $k, h_1, \cdots, h_{s-1}$ for which

$$0 \geqq k \geqq -\max(0, h_1, h_1+h_2, \cdots, h_1+\cdots+h_{\alpha-1}),$$

$$0 \geqq k \geqq -n.$$

Since $c(j,k,l) = 0$ if $k < |l-j|$ we may assume that

$$|h_\beta| \leqq j_\beta \qquad \beta = 1,\cdots,\alpha.$$

If n is large the condition $0 \geqq k \geqq -n$ is implied by the other conditions. Using Lemma 11a we see that

$$\lim_{n\to\infty} \sum_{Q(n)} c(k_1,j_1,k_2) \cdots c(k_\alpha,j_\alpha,k_1) = \sum_{k,h_1,\cdots,h_{\alpha-1}} E(h_1,j_1) \cdots E(h_\alpha,j_\alpha).$$

Note $h_1 + \cdots + h_\alpha = 0$ and thus

$$\sum_{k, h_1, \cdots, h_{\alpha-1}} E(h_1, j_1) \cdots E(h_\alpha, j_\alpha)$$
$$= \sum_{h_1 + \cdots + h_\alpha = 0} E(h_1, j_1) \cdots E(h_\alpha, j_\alpha) \max(0, h_1, \cdots, h_1 + \cdots + h_{\alpha-1}).$$

It follows that

$$\lim_{n\to\infty} [\operatorname{tr}[A_n(f)^\alpha] - \int_{-1}^{1} f(x)^\alpha \{ \sum_0^n p(k,x)^2 \} \Omega(dx)]$$
$$\begin{aligned} &= - \sum_{j_1, \cdots, j_\alpha = 0}^{s} f^\wedge(j_1) \cdots f^\wedge(j_\alpha) \\ &\quad \cdot \sum_{h_1 + \cdots + h_\alpha = 0} E(h_1, j_1) \cdots E(h_\alpha, j_\alpha) \max(0, h_1, \cdots, h_1 + \cdots + h_{\alpha-1}) \\ &= - \sum_{h_1 + \cdots + h_\alpha = 0} f^\wedge(h_1) \cdots f^\wedge(h_\alpha) \max(0, h_1, \cdots, h_1 + \cdots + h_{\alpha-1}), \end{aligned} \tag{2}$$

where

$$f(\cos\theta) = \sum_h f^\wedge(h) e^{ih\theta}.$$

If $C(s) = \text{l.u.b.} \| p(j,x) \|_\infty$ for $0 \leqq j \leqq s$ then it follows from Schwarz's inequality that

$$| c(k,j,l) \| \leqq C(s).$$

For $j_1, \cdots, j_\alpha$ fixed between 0 and s the sum

$$\sum_{Q(n)} c(k_1, j_1, k_2) \cdots c(k_\alpha, j_\alpha, k_1)$$

contains at most $\alpha 2^{\alpha-1} s^\alpha$ non-zero terms, and therefore

$$| \operatorname{tr} A_n(f)^\alpha - \int_{-1}^{1} f(x)^\alpha \{ \sum_{k=0}^{n} p(k,x)^2 \} \Omega(dx) | \leqq \alpha 2^{\alpha-1} s^\alpha C(s)^\alpha \{ \sum_{j=1}^{s} | f^\wedge(j) | \}^\alpha.$$

From this we see that if for $|z|$ sufficiently small we multiply (2) by $-z^\alpha/\alpha$ and sum from $\alpha = 1$ to $\alpha = \infty$, and finally use the Kac-Spitzer identity we obtain

$$\lim_{n\to\infty} \{ \log D_n[F] - \log G[n, F] \} = \log E[F \mid F]. \tag{3}$$

See [4; § 5] for a similar argument carried out in complete detail. What we have just done is to apply Kac's argument as adapted in [4] to the $\{p(k,x)\}_0^\infty$.

Comparing (1) and (3) we see that

$$\lim_{n\to\infty} \{ \log G[n, F] - (n+1) \log G[F] \} = 2 \log E[F \mid \Omega_\alpha] \tag{4}$$

uniformly for $|z|$ sufficiently small. We have

$$\frac{\partial}{\partial z}\log G[n,F]\,|_{z=0} = -\int_{-1}^{1}\{\sum_{k=0}^{n} p(k,x)^2\} f(x)\Omega(dx),$$

$$\frac{\partial}{\partial z}\log G[F]\,|_{z=0} = -\tfrac{1}{\pi}\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}} f(x)\,dx,$$

$$\frac{\partial}{\partial z}\,2\log E[F\mid\Omega_\alpha]\,|_{z=0} = -\sum_{k=1}^{\infty} f^\wedge(k)\Omega_\lambda^\wedge(k)k.$$

It follows that

$$\int_{-1}^{1}\{\sum_{k=0}^{n} p(k,x)^2\} f(x)\Omega(dx)$$
$$= \frac{n+1}{\pi}\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}} f(x)\,dx + \sum_{k=1}^{\infty} f^\wedge(k)\Omega_\lambda(k)k + o(1) \text{ as } n\to\infty,$$

as desired.

12. The inversion of matrices by sections. The following result is related to work of Edrei and Szegö [6] and Duffin [5].

Let $\Omega\in\boldsymbol{S}$ etc. and let $c(x)$ be continuous and never 0 for $x\in I$. We set

$$c^\wedge(j,k) = \int_{-1}^{1} c(x)p(j,x)p(k,x)\Omega(dx),$$

$$d^\wedge(j,k) = \int_{-1}^{1} c(x)^{-1}p(j,x)p(k,x)\Omega(dx).$$

It is easy to see that

$$\sum_{k=0}^{\infty} c^\wedge(j,k)d^\wedge(k,l) = \sum_{k=0}^{\infty} d^\wedge(j,k)c^\wedge(k,l) = \delta(j,l)$$

for $j,l=0,1,\cdots$, so that the infinite matrices

$$[c^\wedge(j,k)]_{j,k=0,1,\cdots} \text{ and } [d^\wedge(j,k)]_{j,k=0,1,\cdots}$$

are inverses.

THEOREM 12a. *Let $\Omega\in\boldsymbol{S}$, etc. If $c(x)\in\boldsymbol{R}_I$ there exists an integer n_0 such that if $n\geqq n_0$, $[c^\wedge(j,k)]_{j,k=0,\cdots,n}$ is invertible and if*

$$[c^\wedge(j,k)]^{-1}{}_{j,k=0,\cdots,n} = [d_n{}^\wedge(j,k)]_{j,k=0,\cdots,n}$$

we have

$$\lim_{n\to\infty} d_n{}^\wedge(j,k) = d^\wedge(j,k) \qquad j,k=0,1,\cdots.$$

Proof. By Theorem 6d there exists an integer n_0 and a constant A such that if $f \in \boldsymbol{E}(n)L^2(\Omega)$ then

$$\|f\|_2 \leqq A\,\|\boldsymbol{E}(n)cf\|_2. \tag{1}$$

This implies, since $\boldsymbol{E}(n)L^2(\Omega)$ has finite dimension, that given $g \in L^2(\Omega)$ there exists a unique $f_n \in \boldsymbol{E}(n)L^2(\Omega)$ such that

$$\boldsymbol{E}(n)cf_n = \boldsymbol{E}(n)g \tag{2}$$

and $\|f_n\| \leqq A\,\|\boldsymbol{E}(n)g\|_2 \leqq A\,\|g\|_2$. Consider $F = c^{-1}g$. We have

$$\begin{aligned} \boldsymbol{E}(n)c[\boldsymbol{E}(n)F] &= \boldsymbol{E}(n)c[\boldsymbol{E}(n)-I]F + \boldsymbol{E}(n)cF \\ \boldsymbol{E}(n)c[\boldsymbol{E}(n)F] &= \boldsymbol{E}(n)g - \boldsymbol{E}(n)c[I-\boldsymbol{E}(n)]F \end{aligned} \tag{3}$$

Subtracting (2) from (3) gives

$$\boldsymbol{E}(n)c[\boldsymbol{E}(n)F - f_n] = -\boldsymbol{E}(n)c[I-\boldsymbol{E}(n)]F$$

and thus (if $n \geqq n_0$) Theorem 6d implies that

$$\|\boldsymbol{E}(n)F - f_n\|_2 \leqq A\,\|c[I-\boldsymbol{E}(n)]F\|_2 \leqq A\,\|c\|_\infty \|[I-\boldsymbol{E}(n)\|]F\|_2$$

It follows that

$$\begin{aligned} \|F - f_n\|_2 \leqq \|F - \boldsymbol{E}(n)F\|_2 &+ \|\boldsymbol{E}(n)F - f_n\|_2 \\ &\leqq \{A\,\|c\|_\infty + 1\}\,\|[I-\boldsymbol{E}(n)]F\|_2, \end{aligned}$$

and finally that

$$\lim_{n\to\infty} \|F - f_n\|_2 = 0. \tag{4}$$

See [3] and [12].

It follows already from (1) that $[c^\wedge(j,k)]_{j,k=0,\cdots,n}$ is invertible if $n \geqq n_0$. We have

$$\sum_{k=0}^{n} c^\wedge(j,k)\,d_n^\wedge(k,l) = \delta(j,l) \qquad j = 0, 1, \cdots, n$$

which is equivalent to

$$\boldsymbol{E}(n)c\Big[\sum_{k=0}^{n} d_n^\wedge(k,l)\,p(k,x)\Big] = p(l,x).$$

It follows from (4) that

$$\lim_{n\to\infty} \Big\|\sum_{k=0}^{n} d_n^\wedge(k,l)\,p(k,x) - p(l,x)\,c(x)^{-1}\Big\|_2 = 0$$

This implies in particular that

$$\lim_{n\to\infty} d_n{}^{\wedge}(k,l) = d^{\wedge}(k,l)$$

as asserted.

Note that in Theorem 12a $c(x)$ is permitted to have complex values while in [5] and [6] only real valued c's are permitted.

REFERENCES.

[1] Glen Baxter, "Polynomials defined by a difference system," *Journal of Mathematical Analysis and Applications*, vol. 2 (1961), pp. 223-263.

[2] ———, "A convergence equivalence related to polynomials on the unit circle," *Transactions of the American Mathematical Society*, vol. 99 (1961), pp. 471-487.

[3] ———, "A norm inequality for a 'finite section' Wiener-Hopf equation," *Illinois Journal of Mathematics*, vol. 7 (1963), pp. 97-103.

[4] Jeffrey Davis and I. I. Hirschman, Jr., "Toeplitz forms and ultraspherical polynomials," to appear in the *Pacific Journal of Mathematics.*

[5] R. Duffin, "The reciprocal of a Fourier Series," *Proceedings of the American Mathematical Society*, vol. 13 (1962), pp. 965-970.

[6] A. Edrei and G. Szegö, "A note on the reciprocal of a Fourier series," *Proceedings of the American Mathematical Society*, vol. 4 (1953), pp. 323-329.

[7] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*, University of California Press, 1958.

[8] I. I. Hirschman, Jr., "Variation diminishing transformations and general orthogonal polynomials," *Canadian Journal of Mathematics*, vol. 16 (1964), pp. 98-107.

[9] ———, "Finite sections of Wiener-Hopf equations and Szegö polynomials," *Journal of Mathematical Analysis and Applications*, vol. 11 (1965), pp. 290-320.

[10] ———, "On a theorem of Szegö, Kac, and Baxter," *Journal d'Analyse Mathématique*, vol. 14 (1965), pp. 225-234.

[11] M. Kac, "Toeplitz matrices, translation kernels, and a related problem in probability theory," *Duke Mathematical Journal*, vol. 21 (1954), pp. 501-509.

[12] E. Reich, "On non-Hermitian Toeplitz matrices," *Mathematica Scandinava*, vol. 10 (1962), pp. 145-152.

[13] G. Szegö, *Orthogonal polynomials*, American Mathematical Society Colloquium Publications XXIII, New York.

[14] ———, "Ein Grenzwertsatz über die Toeplitzschen Determinantan einer reelen positiven Funktion," *Mathematische Annalen*, vol. 76 (1915), pp. 490-503.

[15] J. G. Wendel, "Spitzer's formula: a short proof," *Proceedings of the American Mathematical Society*, vol. 9 (1958), pp. 905-908.

MODIFICATIONS OF SOLVMANIFOLDS AND *G*-INDUCED FLOWS.

By LOUIS AUSLANDER.[1]

Introduction. This paper presents results that complete the investigations begun in [1] and continued in [3]. We will, therefore, use the terminology established in [1] and [3]. However, in order to make the introduction, at least, generally understandable let us recall certain basic definitions.

Let G be a connected, simply connected, solvable Lie group and let C be a closed subgroup of G such that G/C is compact. The compact homogeneous space of right cosets G/C will be called a solvmanifold. In the special case where G is nilpotent, we will call the homogeneous space a nilmanifold. Since we consider G/C as a right coset space, G acts on G/C by left multiplication. If $g(t)$ is a one parameter subgroup of G, then $g(t)$ acts on G/C as a subgroup of G and thus determines a flow on G/C. We will call such flows G-induced flows. Following general conventions, we will call a G-induced flow minimal if every orbit is dense. Now on a solvmanifold G/C, the Haar measure on G induces a measure μ on G/C which is invariant under the action of G. When we speak of G-induced flows on G/C being ergodic, it will always be understood that the measure is μ.

We will need one further concept from [3]. Let $L(G)$ denote the Lie algebra of G and let $X \in L(G)$. Let

$$L(G) = \sum_{i=0}^{r} L_i$$

be the primary decomposition of L with respect to $\operatorname{ad}(X)$. Let $L_a, \cdots, L_r$ be the summands corresponding to eigenvalues that are neither zero nor pure imaginary. Let $\dot{U}(X)$ be the subalgebra of $L(G)$ generated by

$$\sum_{i=a}^{r} L_i.$$

Received May 1, 1965.

Revised April 27, 1966.

[1] During the time that various sections of this paper were in preparation the author received support from the National Science Foundation and the U. S. Naval Research Laboratory.

Then in [3], $U(X)$ was shown to be an ideal in $L(G)$ and called the unstable ideal of X. If $\exp(tX)=g(t)$, we will let

$$\exp(U(X))=U(g(t))$$

and call $U(g(t))$ the unstable subgroup of $g(t)$. *Of course,* $U(g(t))$ *is a normal subgroup of* G. We will call G class R if for any $X\in L(G)$, $U(X)=0$.

With these conventions established, we may relate our work to that in [1] and [3]. The work in [1] concerned itself with minimal G-induced flows on nilmanifolds and ergodic flows on solvmanifolds. The second topic was taken up again in [3]. *In this paper we will complete the work in* [1] *and* [3] *by showing precisely which G-induced flows on solvmanifolds are ergodic and which are minimal.*

The basic development of G-induced flows on solvmanifolds has been along two lines. The first involved the use of unitary representation theory of groups and the Mautner Lemma. The second approach involved detailed information about the structure of compact solvmanifolds. The first approach is used in Chapters V and VII of [1] and the second in Chapter VI of [1]. Both these tools were sharpened in [3] and for the first time certain classes of G-induced flows on general solvmanifolds could be studied. The essentially new ingrediant that this paper adds is the theory of modifications of solvmanifolds as discussed in § 2. Using this we can prove the following results:

Let us begin by recalling the theorem below from § 3 of [3].

THEOREM A. *Let* $g(t)$ *be a G-induced flow on* G/C *and let* $U(g(t))$ *be the unstable subgroup of* $g(t)$. *Then* $g(t)$ *is ergodic on* G/C *if and only if it is ergodic on* $G/(CU(g(t)))^{-}$, *where* $CU(g(t))$ *is the subgroup of* G *generated by* C *and* $U(g(t))$, *and the bar denotes the closure operation.*

One of the main results of this paper is the following:

THEOREM B. *Let* $g(t)$ *be an ergodic G-induced flow on* G/C *with unstable subgroup* $U(g(t))$. *Let* H *be the maximal analytic subgroup of* $(CU(g(t)))^{-}$ *normal in* G. *Then* G/H *is class* R.

Theorem B enables us to apply § 4 of [3] for any ergodic G-induced flow on a solvmanifold to complete the classification problem.

Our second result is the following:

THEOREM C. *Let* M *be a compact solvmanifold with a minimal G-induced flow. Then* M *is homeomorphic to a nilmanifold* N/M *and the G-induced flow is equivalent to a flow that comes from a one parameter subgroup of* N.

1. Modifications of solvmanifolds. Triples $(G, C, g(t))$, where G is a connected, simply connected solvable Lie group, C is a closed subgroup such that G/C is compact, and $g(t)$ is a one parameter subgroup of G will play an important role in our work and will be called admissible. We will want to place three additional sorts of restrictions on such triples.

R_1. C contains no analytic subgroup normal in G.

R_2. If M is the maximal nilpotent normal subgroup of G and CM is closed in G, then $g(t)$ acts minimally on G/CM.

R_3. $g(t)$ acts minimally on G/C.

Let $(G, C, g(t))$ be an admissible triple. We will discuss three types of admissible triples associated with $(G, C, g(t))$. We will call these changes modifications of Type I, II and III.

Type I. *Let* $(G, C, g(t))$ be an admissible triple. Let $C^* \subset G$ be such that either $C^* \supset C$ or $C^* \subset C$ and G/C^* is compact. Then $(G, C^*, g(t))$ is an admissible triple which has properties R_1, R_2 or R_3 if and only if $(G, C, g(t))$ has this property.

Type II. Let $(G, C, g(t))$ be an admissible triple and let H be a connected normal subgroup of G such that $H/H \cap C$ is compact. Then if $g^*(t)$ is the image of $g(t)$ in G/H, $(G/H, CH/H, g^*(t))$ is an admissible triple. It is easy to see that if $(G, C, g(t))$ has property R_2 then so does $(G/H, CH/H, g^*(t))$. The assertion that $(G, C, g(t))$ has property R_3 then $(G/H, CH/H, g^*(t))$ has property R_3 follows from the fact that G/C is a fiber bundle over G/HC with fiber $H/H \cap C$ and $g(t)$ induces a group of fiber bundle mappings of G/C with the property that the induced map on G/CH is exactly $g^*(t)$.

Type III. Let $(G, C, g(t))$ be an admissible triple satisfying property R_1. Then in Chapter VI of [1], Theorem 4.2, we showed how to associate another admissible triple with $(G, C, g(t))$. We will need the following slightly stronger version of that result in this paper.

LEMMA 1. *Let S and G be connected, simply connected solvable Lie groups and let T^* be a compact abelian group of semi-simple automorphisms of S. In addition, let*

1. *$G \subset T^* \cdot S$, where the dot denotes the semi-direct product.*
2. *$C \subset G \cap S$ is such that G/C and S/C are compact.*
3. *G and S generate $T^* \cdot S$.*

4. *If* $p\colon T^*\cdot S \to S$ *denotes the projection mapping then* p *restricted to* G, p/G, *induces a homeomorphism of* G *onto* S.

Then there exists a one parameter subgroup $s(t) \subset S$ *such that*

a) $(S, C, s(t))$ *is an admissible triple satisfying* R_2 *and* R_3 *if and only if* $(G, C, g(t))$ *satisfies* R_2 *and* R_3.

b) S/C *is homeomorphic to* G/C.

c) $s(t)$ *is equivalent to* $g(t)$.

Proof. Condition (b) is clearly true because of 2 and 4 above. The remaining assertions can be seen as follows: Let $L(\ )$ denote the Lie algebra of the group in the bracket. Then $L(T^*\cdot S) = L(T^*) + L(S)$ as a vector space direct sum. Let $X \in L(G)$ determine $g(t)$. Then $X = A + B$, $A \in L(T^*)$ and $B \in L(S)$. Let M denote the maximal nilpotent normal subgroup of G. Then $M/M \cap C$ is compact. Hence by a modification of Type II we see that the image of $g(t)$ in G/MC is dense. Let q denote the projection of $T^*\cdot S \to T^*$. Then by 3 $q(G)$ is dense in T^*. Hence $q(g(t))$ is in general position in T^*. Therefore, by choosing a new semi-direct product representation of $T^*\cdot S$, if necesssary, we may assume that

$$[A, B] = 0$$

and, indeed, that $[L(T^*), B] = 0$. Assume our original semi-direct product was so chosen that the above conditions are satisfied. Let $s(t)$ denote the one parameter subgroup of S determined by B and $\tau(t) \subset T^*$ be the one parameter subgroup of T^* determined by A. Then $g(t) = \tau(t)s(t)$ and $\tau^* s(t) = s(t)\tau^*$ for all $\tau^* \in T^*$ and t. Hence for $g \in G$, $g = \tau^* s$, $s \in S$ we have that

$$\tau(t)s(t)g = \tau(t)s(t)\tau^* s = \tau(t)\tau^* s(t)s.$$

Hence $g(t)$ acting on G/C is equivalent to $s(t)$ acting on S/C relative to the homeomorphism $p^*\colon G/C \to S/C$ induced by p. More explicitly we have that

$$\begin{tikzcd}
G/C \arrow[r, "p^*"] \arrow[d, "g(t)"'] & S/C \arrow[d, "s(t)"] \\
G/C \arrow[r, "p^*"] & S/C
\end{tikzcd}$$

is a commutative diagram.

Let $(G, C, g(t))$ be an admissible triple satisfying property R_1. Let $(S, C, s(t))$ be as in Lemma 1. Then $(S, C, s(t))$ will be called a modification of $(G, C, g(t))$ of Type III.

Definition. Let $(G, C, g(t))$ be an admissible triple and let $(G^{\#}, C^{\#}, g^{\#}(t))$ be another. We will call $(G^{\#}, C^{\#}, g^{\#}(t))$ a modification of $(G, C, g(t))$ if there exists a chain of modifications of Type I, II or III starting with $(G, C, g(t))$ and ending with $(G^{\#}, C^{\#}, g^{\#}(t))$. Note that if $(G^{\#}, C^{\#}, g^{\#}(t))$ is a modification of $(G, C, g(t))$ then each has property R_2 or R_3 if and only if the other one has property R_2 or R_3.

2. Reduction to discrete isotropy group. Let $(G, C, g(t))$ be an admissible triple with property R_1 and let H denote the nil-radical, the maximal nilpotent analytic subgroup, of G. Then by [5], if C_0 denotes the identity component of C, then $C_0 \subset H$ and $H/H \cap C$ is compact. Now by Malcev [4], since H is nilpotent and $H/H \cap C$ is compact, the identity component of $H \cap C$, which is C_0, is normal in H.

Let $\mathrm{ad}_H(C)$ denote the matrix representation of C obtained by letting C on the Lie algebra of H by means of inner automorphisms of G. Let $A(\mathrm{ad}_H(C))$ denote the algebraic hull of $\mathrm{ad}_H(C)$ and $A(\mathrm{ad}_H(C))_0$ the identity component of $A(\mathrm{ad}_H(C))$. Further, let C_1 be the subgroup of C whose image under the mapping ad_H is in $A(\mathrm{ad}_H(C))_0$. Then, since $A(\mathrm{ad}_H(C))_0$ is of finite index in $A(\mathrm{ad}_H(C))$, C_1 is of finite index in C. We thus see that $(G, C_1, g(t))$ is a modification of $(G, C, g(t))$.

We now need Theorems 1, 2 and B of [2] in a stronger form than proven in [2]. In [6] proofs of the necessary generalizations can be found. We will, therefore, satisfy ourselves with stating the stronger theorems we will need.

Let us return to the above situation and notation. Then by an analogous discussion to that in [2] we have

1. There exists a unique maximal nilpotent subgroup M of C_1 containing $[C_1, C_1]C_0$.

2. If θ is any eigenvalue of $\mathrm{ad}_H(C_1)$ then

$$\theta\,|\,\theta\,|^{-1} = \cos 2\pi\zeta + i \sin 2\pi\zeta$$

where $|\,\theta\,|$ denotes the norm of θ and ζ is either 0 or irrational.

THEOREM 1. *Given C_1 as above there exists a connected, simply connected solvable Lie group $S(C_1)$ such that*

1. *$S(C_1) \supset C_1$ and $S(C_1)/C_1$ is compact.*

2. *Let N be the nil-radical of $S(C_1)$. Then $C \cap N = N$ and N/M is compact. Further $N \supset H$.*

3. *Every automorphism of* C_1 *is uniquely extendable to an automorphism of* $S(C_1)$.

4. $C_0 \subset C_1$ *is normal in* $S(C_1)$.

Part 4 is the only part of Theorem 1 not contained in [6]. It may be seen as follows: C_0 is normal in C and in H. Hence the Lie algebra of C_0 is invariant under $A(\mathrm{ad}_H(C_1H))$. But $\mathrm{ad}_H S(C_1) \subset A(\mathrm{ad}_H(C_1H))$ by construction. This proves that C is normal in $S(C_1)$.

The next theorem is the analogue of Theorem B of [2]. Again a proof of this result is contained in [6].

THEOREM 2. *There exists a compact abelian group* T^* *of semi-simple automorphisms of* $S(C_1)$ *such that* $G \subset T^* \cdot S(C_1)$ *as a subgroup, where the dot denotes the semi-direct product. Let* p *denote the projection of* $T^* \cdot S(C_1)$ *onto* $S(C_1)$. *Then* p *maps* G *homeomorphically onto* $S(C_1)$ *and* $C_1 \subset S(C_1) \cap G$.

Theorems 1 and 2 assure us that the admissible triple $(G, C_1, g(t))$ can be modified to $(S(C_1), C_1, s(t))$ by a modification of Type III. But by part 4 of Theorem 1, C_0 is normal in $S(C_1)$ and so if $S = S(C_1)/C_0$, $\Gamma = C_1/C_0$, and $s_1(t)$ denotes the image of $s(t)$ in S_1, we have that $(S, \Gamma, s_1(t))$ is a modification of $(S(C_1), C_1, s(t))$ by a modification of Type II. We will state our result formally in the following theorem.

THEOREM 3. *Let* $(G, C, g(t))$ *be an admissible triple with property* R_1. *Then there exists a modification* $(S, \Gamma, p(t))$ *such that*

1. Γ *is a discrete subgroup of* S.

2. Γ *is isomorphic to a subgroup of finite index of the fundamental group of* G/C.

3. $p(t)$ *acting on* S/Γ *is equivalent to* $g(t)$ *acting on* G/C_1, *where* C_1 *is as above.*

3. Discrete isotropy group. All admissible triples in Sections 3 and 4 will be assumed to have property R_3. Now, let $(S, \Gamma, s(t))$ be an admissible triple with Γ discrete. Let H be the nil-radical of S, i.e. the maximal analytic nilpotent normal subgroup of S. By [4] and [5] if $[H, H]$ denotes the commutator subgroup of H we may apply a modification of Type II and obtain an admissible triple where $[H, H]$ is trivial. To avoid introducing still more notation, we will assume henceforth that this is the case.

We will now discuss the structure of admissible triples $(S, \Gamma, s(t))$ where:

1. Γ is discrete

2. nil-radical is abelian

3. $\mathrm{ad}_H(\Gamma)$ restricted to any invariant subspace modulo real scalar matrices contains no elements of finite order.

We will want still another property to hold which we will now explain. Let $\gamma \in \Gamma$ and let

$$(\mathrm{ad}(\gamma) - I)H = H_1.$$

Then $(\mathrm{ad}(\gamma) - I)$ maps $\Gamma \cap H$ into itself and since H has a basis consisting of elements of $\Gamma \cap H$, so does H_1. Hence $H_1/\Gamma \cap H_1$ is compact. We may define H_k inductively by

$$H_k = (\mathrm{ad}(\gamma) - I)H_{k-1}.$$

Clearly $H_k = H_{k+n} n \geqq 0$ and same k. Further $H_k \cap \Gamma$ is a lattice in H_k. Call such an H_k an admissible subspace. Since $\mathrm{ad}(S)$ is abelian, H_k is invariant under $\mathrm{ad}(S)$ or normal in S. Hence we may apply a modification of Type II to S with respect to H_k. Clearly by a succession of modifications of Type II we can find a modification $(S^*, I^*, s^*(t))$ such that S^* has exactly one non-trivial admissible subspace. We may, therefore, consider the further hypothesis

4. S has exactly one non-trivial admissible subspace. Call it A.

Lemma 2. *Let* $S, A, \Gamma, s(t)$ *satisfy* 1-4. *Then*

$$1 \to A \to S \to S/A \to 1$$

is a split extension and S/A *is nilpotent.*

Proof. We will prove this lemma by induction on $\dim. H - \dim. A$. If it is zero, let $\gamma \in \Gamma$ be such that $\mathrm{ad}(\gamma)$ is in general position. Let V be the eigenvalue one subspace of $\mathrm{ad}(\gamma)$, clearly since $\mathrm{ad}(\gamma)$ is an automorphism of $L(S)$, if $X, Y \in L(S)$, then $[X, Y]$ is an eigenvalue one vector in $L(S)$ also. Hence V is a subalgebra. But since S/A is abelian, we must have since $A = H$ that the exact sequence splits.

Now let $\dim. H - \dim. A > 0$. Let $V \subset H$ be the eigenvalue one subspace of $\mathrm{ad}(s)$ all $s \in S$. V_1 is non-empty since $\mathrm{ad}(S)$ is abelian. Clearly V_1 is a normal subgroup of S. Dividing by V_1, we get a new $S_1 \supset A$ which by induction splits. Since $V_1 \cap A = 0$ this proves that $1 \to A \to S \to S/A \to 1$ also splits.

Lemma 3. *Let* S *and* Γ *satisfy* 1-4. *Then* $\Gamma^* \supset \Gamma$ *exists such that* Γ^*/Γ *is finite and* $A/\Gamma^* \cap A$ *is compact. Further we can find a direct sum representation for* $S = N \cdot A$ *such that* $N/\Gamma^* \cap N$ *is compact, where* $N \cong S/A$.

Proof. Let $L = A \cap \Gamma$. Then A/L is compact. Let $\gamma_0 \in \Gamma$ be such that $(\mathrm{ad}(\gamma_0) - I)$ restricted to A is nonsingular. Let B be the restriction of $\mathrm{ad}(\gamma_0)$ to A and let Δ be the determinant of $\mathrm{ad}(\gamma_0) - I)$. Then Δ is an integer. Since $B - I$ is nonsingular we may choose a semi-direct product representation of S so that $\gamma_0 \in N$. Let $\gamma \in \Gamma$ then $\gamma = an$, where $a \in A$ and $n \in N$. Now the commutator $[\gamma_0, \gamma]$ gives $(B - I)a \in L$. This can be seen as follows: Let $\delta \in [\Gamma A/A, \Gamma A/A]$. Then δ is central in N. Let δ have the pre-image $\gamma^* = a(\delta^*)n(\delta^*)$ in Γ. Then $[\gamma, \gamma^*] \in L$ because $[\gamma_0, \gamma^*] \equiv 0 \bmod L$. Hence $(B - I)a(\delta^*) \in L$. Since for arbitrary γ, we have $[\gamma_0, \gamma] \equiv \delta \bmod L$, we have

$$(B - I)a \equiv (B - I)a(\delta) \bmod L.$$

Hence $(B - I)a(\gamma) \in L$ for all $\gamma \in \Gamma$. Hence if we enlarge Γ to Γ^* by including $L \otimes 1/\Delta$, we get

$$\begin{tikzcd}
1 \arrow[r] & L \otimes \Delta^{-1} \arrow[r] \arrow[d] & \Gamma^* \arrow[r] \arrow[d] & \Gamma/L \arrow[r] \arrow[d] & 1 \\
1 \arrow[r] & A \arrow[r] & S \arrow[r] & N \arrow[r] & 1
\end{tikzcd}$$

such that $S = N \cdot A$ and $N/N \cap \Gamma^*$ is compact.

This shows that $[N, N] \cap \Gamma^*$ is such that $[N, N]/[N, N] \cap \Gamma^*$ is compact. Therefore, we can apply a modification of Type II to get the following condition:

5. The nil-radical of S is the admissible subspace $A, S = R^s \cdot A$ and $R^s/\Gamma \cap R^s$ is compact.

THEOREM 4. *Let $(G, C, g(t))$ be an admissible triple. Then G/C is finitely covered by a nil-manifold and $g(t)$ is equivalent to a nil-flow on this covering space.*

Proof. We will actually see that $(S, \Gamma, s(t))$ constructed in § 2 is such that S is nilpotent. This will be an easy consequence of the results in this section if we can show that conditions 1-5 imply that A is trivial. But it is a straightforward computation to show that under conditions 1-5 for any one parameter group in general position there is an orbit whose closure is an s dimensional torus. This means that A must be trivial and that S is nilpotent and we have our theorem.

4. Theorem C.

THEOREM C. *Let $(G, C, g(t))$ be admissible. Then G/C is a nil-manifold and $g(t)$ is equivalent to a nil-flow.*

Proof. By Theorem 4, $S(C_1)$ in Theorem 1 may be assumed to be nilpotent and to give this point stronger emphasis we will change notation and call $S(C_1), N(C_1)$, and denote the admissible triple by $(N(C_1), C_1, n(t))$. Then $[N(C_1), N(C_1)]/C_1 \cap [N(C_1), N(C_1)]$ is compact and so we may apply a modification of Type II to find an admissible triple $(N, C^*, n^*(t))$ where N is abelian.

Now $G \subset T^* \cdot N(C_1)$ and if M is the nil-radical of G then $T^*N(C_1)/M$ is abelian. Hence $M \supset [N(C_1), N(C_1)]$ and we may consider

$$G^* = G/[N(C_1), N(C_1)]$$

with uniform subgroup $CM/M = C^*$. It is easily seen that G^*/C^* is a compact locally euclidean manifold. But by the structure theory of solvmanifolds which are also locally euclidean manifolds as developed in [6] or [7], if M^* is the nil-radical of G^*, $G^* = R^s \cdot M^*$, where M^* is abelian and R^s acts on M^* as a compact abelian group of automorphisms. Since $(G_1^*, C^*, n^*(t))$ is admissible $n^*(t)$ is in general position in the compact abelian group of automorphisms determined by R^s. Now C^* is non-abelian. Hence there exists in M^* a subspace A which is admissible. We may alter G^* by modifications of Type II and III until $M^* = A$. We may now argue exactly as in § 3 to conclude that A must be zero and so C^* is abelian and we are done.

5. Theorem B. *All admissible triples in this section will be assumed to have property* R_2.

THEOREM D. *Let* $(G, C, g(t))$ *be an admissible triple with property* R_1 *and let the unstable subgroup of* $g(t)$ *be trivial. Then* G *is class* R.

We will begin by showing that Theorem D implies Theorem B. Let $g(t)$ act ergodically in G/C and let ψ be the natural homomorphism

$$\psi\colon G \to G/H = G^*$$

where H is the nil-radical of G. Further, let $(\psi(CU(g(t))))^- = C^*$ and $\psi(g(t)) = g^*(t)$. Then it is obvious that the admissible triple $(G^*, C^*, g^*(t))$ has property R_2. But then D implies B.

We will now see that Theorem 3 and the following theorem implies Theorem D.

THEOREM D′. *Let* $(S, \Gamma, p(t))$ *be an admissible triple with* Γ *discrete and such that the unstable subgroup of* $p(t)$ *is trivial. Then* S *is class* R.

So let us assume that Theorem D′ holds and verify Theorem D. Let all notations be as above. By Theorem 4.2 of [2], Γ has a subgroup of finite index which is nilpotent. Hence $(G, C, g(t))$ has the property of having its fundamental group contain a nilpotent subgroup of finite index. But then we may apply Appendix A of [2] to conclude that G is class R.

Before beginning the proof of Theorem D′, let us quote the following standard theorem which permits us to make one further reduction in our problem.

THEOREM 5. *Let* $(S, \Gamma, p(t))$ *be an admissible triple with* Γ *discrete and such that* $U(p(t))$ *is trivial. Let* N *be the nil-radical of* S *and let* $\eta\colon S \rightarrow S/[N,N]$ *be the natural homomorphism. Then if* $S^* = S/[N,N]$, $\Gamma^* = \eta(\Gamma)$ *and* $p^* = \eta(p(t))$, *we have that* $(S^*, \Gamma^*, p^*(t))$ *is an admissible triple. Further,* S *is class* R *if and only if* S^* *is class* R.

Thus we see that Theorem D′ is true in general if and only if it is true in the special case where the nil-radical of S is abelian. We will now prove Theorem D′ with this additional assumption.

We have, therefore, the commutative diagram

$$\begin{tikzcd}
1 \arrow[r] & Z^a \arrow[r] \arrow[d] & \Gamma \arrow[r] \arrow[d] & Z^b \arrow[r] \arrow[d] & 1 \\
1 \arrow[r] & Z^a \otimes R \arrow[r] & S \arrow[r] & Z^b \otimes R \arrow[r] & 1
\end{tikzcd}$$

where Z^n denotes the integers taken n times, R denotes the reals and $\alpha(z) = z \otimes 1$ for $z \in Z^a$ or Z^b. Further $S/\beta(\Gamma)$ is compact, $\beta(\Gamma)$ is discrete in S, and $j(p(t))$ has dense orbit in $Z^b \otimes R/\alpha(Z^b)$. Now each $w \in Z^b \otimes R$ induces a linear transformation, $\mathrm{ad}(w), Z^a \otimes C$, where C is the complex numbers essentially by inner automorphisms of S. Let

$$\mathrm{ad}(w) = n(w) \cdot s(w)$$

be the decomposition of $\mathrm{ad}(w)$ into its semi-simple and unipotent parts. Since $Z^b \otimes R$ is abelian, the mapping

$$s\colon w = s(w) \qquad w \in Z^b \otimes R$$

is a homomorphism of $Z^b \otimes R$ into C^a, where C^a denotes a copies of the *multiplicative group* of the complex field. Further, since the unstable ideal of $p(t)$ is trivial $s(p(t))$ is contained in $(C^*)^a$, where C^* denotes the multiplicative group of complex numbers of norm one.

We will now show that there exists at least one $z \in Z^b$ such that $s(\alpha(z)) \in (C^*)^a$. Note first that $\mathrm{ad}(\alpha(z)), z \in Z^b$, is an integer unimodular matrix relative to the basis $\alpha(z_1), \cdots, \alpha(z_a)$ of $Z^a \otimes C$, where $z_1, \cdots, z_a$ are generators of Z^a. Since an integer unimodular matrix all of whose eigenvalues have norm sufficiently close to one has the property that its eigenvalues all have norm one, and since $p(t)$ gets arbitrarily close to some element of $\alpha(Z^b)$, we have our first assertion.

Now let $z_1 \in Z^b$ be such that $s(\alpha(z_1)) \in (C^*)^a$. Let $z_2, \cdots, z_b$ be such that $z_1, \cdots, z_b$ generate a subgroup of finite index in Z^b. Let

$$S_1 = j^{-1}(\alpha(Z^{b-1}) \otimes R)$$

where Z^{b-1} has $z_2, \cdots, z_b$ as generators, and let

$$\Gamma_1 = S_1 \cap \beta(\Gamma).$$

Then S_1/Γ_1 is compact. Let

$$Z^b \otimes R = (\alpha(z_1))^n \otimes R \oplus \alpha(Z^{b-1}) \otimes R, \quad n = 0, \pm 1, \cdots.$$

Then let $p(t) = p_0(t) \oplus p_1(t)$ be the corresponding direct sum decomposition of $p(t)$. Note that $p_1(t)S_1$ satisfies the condition that $s(p_1(t)) \in (C^*)^a$ since $s((\alpha(z_1))^n \otimes (C^*)^a$. Hence $(S_1, \Gamma_1, p(t))$ satisfies all the hypotheses of our theorem. Thus an elementary induction proves Theorem D′ and hence Theorem D.

REFERENCES.

[1] L. Auslander, L. Green and F. Hahn, *Flows on Homogeneous Spaces*, Annals of Mathematics Study No. 53.

[2] L. Auslander, "Discrete uniform subgroups of solvable Lie groups," *Transactions of the American Mathematical Society*, vol. 99 (1961), pp. 398-402.

[3] L. Auslander and L. Green, "G-induced flows," *American Journal of Mathematics*, vol. 88 (1966), pp. 43-60.

[4] A. Malcev, "On a class of homogeneous spaces," *American Mathematical Society Translations*, Series 1, No. 39 (1951).

[5] G. D. Mostow, "Factor spaces of solvable Lie groups," *Annals of Mathematics*, vol. 60 (1954), pp. 1-27.

[6] C. Perry, *Structure of Compact Solvmanifolds*, Thesis, Purdue University (1965).

[7] L. Auslander and M. Auslander, "Solvable Lie groups and locally Euclidean Riemann spaces," *Proceedings of the American Mathematical Society*, vol. — (1958), pp. 933-941.

SPECTRAL SYNTHESIS PROBLEMS FOR INVARIANT SUBSPACES ON GROUPS. I.

By John E. Gilbert.

1. In this paper we consider the problem of characterizing the closed translation invariant subspaces of the space $\mathcal{C}(G)$ of continuous complex-valued functions on a locally compact abelian group G. The topology on $\mathcal{C}(G)$ is that of compact convergence. Spectral synthesis is said to hold for a set Δ of exponential polynomials on G (see Definition 3.1) if there is one and only one closed translation invariant subspace τ of $\mathcal{C}(G)$ whose spectrum is Δ. We show that the spectrum of any such subspace τ is dense in τ if τ consists of all functions $f \in \mathcal{C}(G)$ satisfying the convolution equation

$$f * \mu^*(x) = \int_G f(xy)\, d\mu(y) = 0, \quad x \in G,$$

where μ is a given measure with compact support in G. Thus spectral synthesis holds for the subspaces of $\mathcal{C}(G)$ orthogonal to singly-generated, weak*-closed ideals in the algebra $\mathcal{M}_c(G)$ of measures with compact support in G. This extends to arbitrary locally compact abelian groups the result of Malgrange [5], Ehrenpreis [2] for $\boldsymbol{R}^p$ and of Lefranc [3] for $\boldsymbol{Z}^p$. The result is obtained by using techniques of abstract harmonic analysis applied to functions on compactly generated subgroups of G.

The basic ideas for this paper are contained in my D.Phil. thesis (Oxford, 1963) written under the supervision of Professor G. Temple. I should like to thank Dr. H. Reiter for stimulating instruction covering many topics in harmonic analysis.

2. Throughout the paper we shall adopt the notation and terminology current in harmonic analysis (cf. [1], [11]). Locally compact abelian groups are denoted by Roman letters $g, G, G_\alpha, \cdots$ and their respective Pontrjagin character groups by corresponding Greek letters $\gamma, \Gamma, \Gamma_\alpha, \cdots$. German letters $\mathfrak{G}, \mathfrak{G}_\alpha, \cdots$ will then signify the generalized character groups of $G, G_\alpha, \cdots$. Bilinear forms connecting g with γ and $G, G_\alpha, \cdots$ with $\mathfrak{G}, \mathfrak{G}_\alpha, \cdots$ we denote by $\langle x, \sigma \rangle$, (x, λ) respectively, where $x \in g, G, G_\alpha, \cdots, \sigma \in \gamma$ and $\lambda \in \mathfrak{G}, \mathfrak{G}_\alpha, \cdots$.

Received August 6, 1964.
Revised May 5, 1965.

The notation for the p-fold products of classical groups will be (i) real numbers: $\boldsymbol{R}^p$, (ii) rational integers: $\boldsymbol{Z}^p$ and (iii) complex numbers: $\boldsymbol{C}^p$. Except for the classical groups the group operation will be written multiplicatively. The symbols dx, $d\xi$, etc. represent the Haar measure on the groups in question and the Haar measure of any compact group will always be taken as 1.

All functions and measures considered are complex-valued. The translate f_a of any function f on $G, G_\alpha \cdots$ $(a \in G, G_\alpha, \cdots)$ is given by $f_a(x) = f(ax)$.

3. It is well-known that any locally compact abelian group G is the inductive limit of compactly generated groups G_α, where α runs over an index set A (cf. [11] p. 109), and any compact set K in G is contained in one of these groups G_α.

If F_G denotes the torsion subgroup of G, i. e., the subgroup of G of all elements of finite order, the quotient group G/F_G is torsion free. Thus the group of homomorphisms of G/F_G into $\boldsymbol{R}^1$ (the real characters) separates elements of G/F_G ([4] p. 157). We choose once and for all a subset $\{\chi_j\}$ $(j \in J)$ of the real characters such that, for each $x \in G/F_G$, there is a χ_j satisfying $\chi_j(x) \neq 0$ unless x is the identity. The set $\{\chi_j\}$ can be regarded as real characters on G by setting $\chi_j(x) = \chi_j(xF_G)$ since χ_j necessarly vanishes on an element of finite order.

Definition 3.1. Let λ be a fixed element of $\mathfrak{G}$. Then the function $P(\cdot, \lambda)$ defined on G by

$$P(x,\lambda) = [P(\chi_j(x))](x,\lambda), \quad x \in G,$$

where $P(\chi_j(x))$ is a polynomial invoiving only finitely many of the χ_j's $(j \in J)$, is called an EXPONENTIAL POLYNOMIAL.

For the classical cases of $G = \boldsymbol{R}^p$ or $\boldsymbol{Z}^p$ Definition 3.1 agrees with the usual one (see [5], [3]). The restriction to $G_\alpha (\alpha \in A)$ of the fixed set $\{\chi_j\}$ is again a set of real characters having the property that it separates elements in the torsion-free quotient group of G_α with respect to its torsion subgroup. Hence any exponential polynomial on G_α can be taken to be of the form

$$Q(\xi,\lambda') = [Q(\chi_j(\xi))](\xi,\lambda'), \quad \xi \in G_\alpha$$

for fixed $\lambda' \in \mathfrak{G}_\alpha$.

We denote by $\mathcal{E}(G)$ the vector space, under the topology of compact convergence, of all continuous functions on G. By an invariant subspace τ of $\mathcal{E}(G)$ we shall mean a closed subspace invariant under translation by elements of G, i. e., $f_a \in \tau$ when $f \in \tau$ and $a \in G$. Since any exponential polynomial $P(\cdot,\lambda)$ belongs to $\mathcal{E}(G)$, the set of all exponential polynomials in an invariant subspace τ (to be denoted by $sp(\tau)$) is called the SPECTRUM of τ. There are similar definitions for $\mathcal{E}(G_\alpha)$ and for invariant subspaces of $\mathcal{E}(G_\alpha)$.

The space $\mathfrak{M}_C(G)$ $(\mathfrak{M}_C(G_\alpha))$ dual to $\mathcal{E}(G)$ $(\mathcal{E}(G_\alpha))$ consists of Radon measures having compact support in $G(G_\alpha)$. Because any compact set in G is contained in one of the G_α's it is easy to show that $\mathfrak{M}_C(G)$ is the algebraic inductive limit of the spaces $\mathfrak{M}_C(G_\alpha)$ corresponding to the inductive limit of the groups G_α. On any space of measures we put the weak* topology. As a topological algebra under the usual addition and convolution of measures, $\mathfrak{M}_C(G)$ possesses an involution $\mu\to\mu^*$ given by $\mu^*(x)=\mu(x^{-1})$. The weak*-closed subspaces of $\mathfrak{M}_C(G)$ orthogonal to the invariant subspaces of $\mathcal{E}(G)$ are precisely the weak*-closed ideals of $\mathfrak{M}_C(G)$.

Convolution of elements of $\mathcal{E}(G)$ with those in $\mathfrak{M}_C(G)$ exists and belongs to $\mathcal{E}(G)$. For a given $\mu\in\mathfrak{M}_C(G)$ we denote by τ_μ the invariant subspace of $\mathcal{E}(G)$ of functions satisfying $f*\mu^*=0$ on G. Clearly $P(\cdot,\lambda)\in sp(\tau_\mu)$ if and only if

$$P(\cdot,\lambda)*\mu^*(x)=\int_G P(\chi_j(xy))(xy,\lambda)\,d\mu(y)=0, \qquad x\in G.$$

The main result of the paper is the following theorem.

THEOREM 3.2. *Let $\mu\in\mathfrak{M}_C(G)$ be given. Then the spectrum $sp(\tau_\mu)$ of τ_μ is dense in τ_μ. Equivalently—spectral synthesis holds for the invariant subspaces of $\mathcal{E}(G)$ orthogonal to singly-generated, weak*-closed ideals in $\mathfrak{M}_C(G)$.*

This extends to arbitrary locally compact abelian groups the corresponding result for the classical groups (see [5], [2] and [3]). The proof will follow from a series of results. It is sufficient to show that, whenever $\nu\in\mathfrak{M}_C(G)$ satisfies $P(\cdot,\lambda)*\nu^*=0$ on G for all $P(\cdot,\lambda)\in sp(\tau_\mu)$, $f*\nu^*=0$ for each $f\in\tau_\mu$. This criterion is both necessary and sufficient; there is a similar criterion for invariant subspaces of $\mathcal{E}(G_\alpha)$ $(\alpha\in A)$.

Any exponential polynomial $Q'(\cdot,\lambda')=[Q'(\chi_j(\cdot))](\cdot,\lambda')$ on G_α can be extended to an exponential polynomial $Q(\cdot,\lambda)$ on G by setting

$$(1) \qquad Q(x,\lambda)=[Q'(\chi_j(x))](x,\lambda), \quad x\in G,$$

where $\lambda\in\mathfrak{G}$ satisfies $\lambda G_\alpha^\perp=\lambda'$.[1] The following simple lemma will be useful.

LEMMA 3.3. *Let $\mu\in\mathfrak{M}_C(G)$ be given and let $Q'(\cdot,\lambda')$ be an exponential*

---

[1] $G_\alpha^\perp$ is the subgroup of $\mathfrak{G}$ orthogonal to G_α.

polynomial on G_α *for which* $Q'(\cdot,\lambda') * \mu^* = 0$ *on* G_α. *Then, for each fixed* $x \in G$, *the function* $Q'_x(\cdot,\lambda')$ *defined by*

$$Q'_x(\xi,\lambda') = Q'(\chi_j(x\xi))(x\xi,\lambda'), \quad \xi \in G_\alpha,$$

also satisfies $Q'_x(\cdot,\lambda') * \mu^* = 0$ *on* G_α.

Proof. By definition

$$Q'(\cdot,\lambda') * \mu^*(\eta) = \int_{G_\alpha} Q'(\eta\xi,\lambda')\, d\mu(\xi) \qquad \eta \in G_\alpha$$

$$(2) \qquad = \sum_{(m)} a_{(m)} \int_{G_\alpha} \chi_{j_1}{}^{n_1}(\eta\xi) \cdots \chi_{j_k}{}^{n_k}(\eta\xi)\,(\eta\xi,\lambda')\, d\mu(\xi)$$

$$= (\eta,\lambda') \sum_{(m)} \chi_{j_1}{}^{n_1}(\eta) \cdots \chi_{j_k}{}^{n_k}(\eta) \int_{G_\alpha} T_{(m)}(\chi_j(\xi))(\xi,\lambda')\, d\mu(\xi)$$

where $(m) = (n_1, n_2, \cdots, n_k)$ and $\sum_{(m)}$ denotes summation over all such finite arrays (only finitely many $a_{(m)}$ are non-zero). The functions $T_{(m)}$ are polynomials in $\chi_j(\xi)$. Similarly, for each $x \in G$, we have

$$\int_{G_\alpha} Q'_x(\eta\xi,\lambda')\, d\mu(\xi)$$

$$= (x\eta,\lambda') \sum_{(m)} \chi_{j_1}{}^{n_1}(x\eta) \cdots \chi_{j_k}{}^{n_k}(x\eta) \int_{G_\alpha} T_{(m)}(\chi_j(\xi))(\xi,\lambda')\, d\mu(\xi).$$

As G_α is isomorphic to $\boldsymbol{R}^p \times \boldsymbol{Z}^q \times g$ (p, q non-negative integers, g compact) the values of the χ_j's used in $Q'(\cdot,\lambda')$ are determined by their values on $\boldsymbol{R}^p \times \boldsymbol{Z}^q$. We may assume that such $\chi_j \not\equiv 0$ on G_α. Now relation (2) can vanish for all $\eta \in G_\alpha$ only if

$$\int_{G_\alpha} T_{(m)}(\chi_j(\xi))(\xi,\lambda')\, d\mu(\xi)$$

vanishes for each $T_{(m)}$. Hence $Q'_x(\cdot,\lambda') * \mu^* = 0$ on G_α.

THEOREM 3.4. *The statement of Theorem 3.2 is true if it is true for each compactly generated subgroup* $G_\alpha (\alpha \in A)$ *of* G.

Proof. Let $\nu \in \mathfrak{M}_c(G)$ satisfy $P(\cdot,\lambda) * \nu^* = 0$ whenever $P(\cdot,\lambda) \in sp(\tau_\mu)$ and let $f \in \tau_\mu$. The supports of μ, ν are contained in subgroups G_β, $G_\gamma (\beta, \gamma \in A)$ respectively and, by the inductive limit property of G, there is a group G_α which contains both supports. Thus $\mu, \nu \in \mathfrak{M}_c(G_\alpha)$. Now

$$\int_G f(x\eta y)\,d\mu(y)=\int_{G_\alpha} f(x\eta\xi)\,d\mu(\xi)$$
$$=\int_{G_\alpha} f_x(\eta\xi)\,d\mu(\xi)=0,\quad \eta\in G_\alpha, x\in G.$$

Hence, for each $x\in G$, f_x restricted to G_α belongs to the invariant subspace $\tau_\mu{}'$ of $\mathcal{B}(G_\alpha)$ generated by functions satisfying $h*\mu^*=0$ on G_α. If $Q'(\cdot,\lambda')\in sp(\tau_\mu{}')$ is extended to G via formula (1) then

$$(Q(\cdot,\lambda)*\mu^*)(x\eta)=(Q'_x(\cdot,\lambda')*\mu^*)(\eta)\qquad x\in G,\eta\in G_\alpha.$$

By Lemma 3.3 $Q'_x(\cdot,\lambda')*\mu^*=0$ on G_α and so $Q(\cdot,\lambda)\in sp(\tau_\mu)$. Consequently $Q'(\cdot,\lambda')*\nu^*=0$ on G_α for all $Q'(\cdot,\lambda')\in sp(\tau_\mu{}')$. Hence, by hypothesis,

$$\int_{G_\alpha} f(xy)\,d\nu(y)=\int_{G_\alpha} f(x\xi)\,d\nu(\xi)=\int_{G_\alpha} f_x(\xi)\,d\nu(\xi)=0,\qquad x\in G.$$

This shows that $sp(\tau_\mu)$ is dense in τ_μ.

The proof of Theorem 3.2 has, therefore, been reduced to consideration of compactly generated groups. This will be the object of the next section.

4. It is well-known that any compactly generated group is isomorphic to $\mathbf{R}^p\times\mathbf{Z}^q\times g$ where p, q are non-negative integers and g is a compact group. For the proof of Theorem 3.2 we consider first $\mathbf{R}^p$, secondly we pass to $\mathbf{R}^p\times\mathbf{Z}^q$ by a rather simple argument and finally we reach arbitrary compactly generated groups.

A. $G=\mathbf{R}^p$. Essentially this is the result of Malgrange ([5], [6]). For the sake of completeness we reproduce his proof (actually the proof given here departs slightly from that of [6]).

THEOREM 4.1. *Let $\mu,\nu\in\mathfrak{M}_C(\mathbf{R}^p)$ have Fourier-Laplace Transforms*[2] *$M(\lambda)$, $N(\lambda)$ such that $N(\lambda)/M(\lambda)$ is an entire function on $\mathbf{C}^p$. Then there exist sequences $\{\delta_n\}$, $\{\psi_n\}$ in $\mathfrak{M}_C(\mathbf{R}^p)$ for which*

(i) *$\delta_n\in\mathcal{D}(\mathbf{R}^p)$ and, as n tends to infinity, δ_n converges to the Dirac-δ-measure in the space $\mathcal{E}'(\mathbf{R}^p)$ of Schwartz* [9],

(ii) *for all n, $\delta_n*\nu=\mu*\psi_n$.*

Proof. When $N(\lambda)/M(\lambda)$ is entire there corresponds to each polynomial $P(x)$ a $\nu_p\in\mathcal{E}'(\mathbf{R}^p)$ such that $(P\mu)*\nu=\mu*\nu_p$ ([6] p. 20). Convolving both sides (if necessary) with a function $\phi\in\mathcal{D}$ we may assume that

[2] In general the Fourier-Laplace Transform is defined by $M(\lambda)=\int_G(x,\lambda)\,d\mu(x)$.

$(P\mu * \phi)$, $(\nu_p * \phi) \in \mathcal{D}$. Choosing sequences $\{P_n\}$ of polynomials and $\{\phi_n\}$ contained in $\mathcal{D}(\boldsymbol{R}^p)$ with $\{P_n\mu\}$, $\{\phi_n\}$ approximating to the Dirac-δ-measure in $\mathcal{E}'(\boldsymbol{R}^p)$ we see that $(P_n\mu * \phi_n) * \nu = \mu * (\nu_{p_n} * \phi_n)$ for all n. But $(P_n\mu * \phi_n)$, $(\nu_{p_n} * \phi_n) \in \mathfrak{M}_C(\boldsymbol{R}^n)$. Properties (i), (ii) now follow at once.

COROLLARY 4.2. *Let* $\mu \in \mathfrak{M}_C(\boldsymbol{R}^p)$. *Then* $sp(\tau_\mu)$ *is dense in* τ_μ.

Proof. If $\nu \in \mathfrak{M}_C(\boldsymbol{R}^p)$ satisfies $P(\cdot, \lambda) * \nu^* = 0$ for all $P(\cdot, \lambda) \in sp(\tau_\mu)$ then $N(\lambda)/M(\lambda)$ is an entire function (cf. [5] p. 283). Thus, by Theorem 4.1,

$$f * (\mu * \psi_n)^* = (f * \mu^*) * \psi_n^* = (f * \nu^*) * \delta_n^* = 0,$$

whenever $f \in \tau_\mu$. Now $(f * \nu^*) * \delta_n^*$ converges in $\mathcal{D}'(\boldsymbol{R}^p)$ (the notation is that of Schwartz [9]) to $f * \nu^*$. Consequently $f * \nu^* = 0$ on $\boldsymbol{R}^p$ and so $sp(\tau_\mu)$ is dense in τ_μ.

B. $G = \boldsymbol{R}^p \times \boldsymbol{Z}^q$. Since $\boldsymbol{R}^p \times \boldsymbol{Z}^q$ is a subgroup of $\boldsymbol{R}^{p+q}$, any $\mu \in \mathfrak{M}_C(\boldsymbol{R}^p \times \boldsymbol{Z}^q)$ belongs also to $\mathfrak{M}_C(\boldsymbol{R}^{p+q})$. Let τ_μ, σ_μ denote the invariant subspaces of $\mathcal{E}(\boldsymbol{R}^p \times \boldsymbol{Z}^q)$, $\mathcal{E}(\boldsymbol{R}^{p+q})$ generated by functions satisfying $f * \mu^* = 0$ on $\boldsymbol{R}^p \times \boldsymbol{Z}^q$ and $g * \mu^* = 0$ on $\boldsymbol{R}^{p+q}$ respectively. Using Lemma 3.3 we can soon show that $sp(\sigma_\mu)$ consists of exponential polynomials on $\boldsymbol{R}^{p+q}$ whose restrictions to $\boldsymbol{R}^p \times \boldsymbol{Z}^q$ belong to $sp(\tau_\mu)$.

For each neighborhood U_n of a fundamental family of compact symmetric neighborhoods which shrink to the identity in $\boldsymbol{R}^{p+q}$ as n tends to infinity let $u_n(x)$ be a continuous non-negative function with support in U_n such that

$$\text{(i)} \quad u_n(x) = u_n(-x), \qquad \text{(ii)} \quad \int_{\boldsymbol{R}^{p+q}} u_n(x)\, dx = 1.$$

When $f \in \tau_\mu$ we denote by f_n the function defined on $\boldsymbol{R}^{p+q}$ by

$$f_n(x) = \int f(y) u_n(x - y)\, dy, \qquad y \in \boldsymbol{R}^p \times \boldsymbol{Z}^q.$$

LEMMA 4.3. *The function* f_n *is a continuous function on* $\boldsymbol{R}^{p+q}$ *belonging to* σ_μ. *Moreover, for any* $\nu \in \mathfrak{M}_C(\boldsymbol{R}^p \times \boldsymbol{Z}^q)$,

$$(3) \qquad \lim_{n\to\infty} \int_{\boldsymbol{R}^{p+q}} f_n(x + y)\, d\nu(y) = \int_{\boldsymbol{R}^p \times \boldsymbol{Z}^q} f(x + y)\, d\nu(y), \qquad x \in \boldsymbol{R}^p \times \boldsymbol{Z}^q.$$

Proof. (cf. [7] p. 479). Trivially $f_n \in \mathcal{E}(\boldsymbol{R}^{p+q})$ while

$$f_n * \mu^* = (f * \mu^*) * u_n = 0.$$

Thus $f_n \in \sigma_\mu$. Now

$$|(f*\nu^*)(x)-(f_n*\nu^*)(x)|=|(f-f*u_n)*\nu^*(x)|$$
$$\leqq \sup_{y\in K}|(f-f*u_n)(y)|\cdot\int|d\nu|,$$

where K is a compact set in $\boldsymbol{R}^p\times\boldsymbol{Z}^q$ depending on x and the support of ν. Given any $\epsilon>0$ and any compact set K in $\boldsymbol{R}^p\times\boldsymbol{Z}^q$ there is a neighborhood U_n of the identity in $\boldsymbol{R}^{p+q}$ such that

$$\sup_{y\in K}|(f-f*u_n)(y)|<\epsilon$$

(cf. [11] p. 52). This establishes (3).

THEOREM 4.4. *Let* $\mu\in\mathcal{M}_C(\boldsymbol{R}^p\times\boldsymbol{Z}^q)$. *Then* $sp(\tau_\mu)$ *is dense in* τ_μ.

Proof. With the notation of Lemma 4.3, $f_n\in\sigma_\mu$. Consequently (cf. Corollary 4.2) $f_n*\nu^*=0$ on $\boldsymbol{R}^{p+q}$ for all n. Hence

$$(f*\nu^*)(x)=\lim_{n\to\infty}(f_n*\nu^*)(x)=0, \qquad x\in\boldsymbol{R}^p\times\boldsymbol{Z}^q,$$

when $\nu\in\mathcal{M}_C(\boldsymbol{R}^p\times\boldsymbol{Z}^q)$ satisfies $P(\cdot,\lambda')*\nu^*=0$ on $\boldsymbol{R}^p\times\boldsymbol{Z}^q$ for any $P(\cdot,\lambda')\in sp(\tau_\mu)$. Thus $sp(\tau_\mu)$ is dense in τ_μ.

C. $G=\boldsymbol{R}^p\times\boldsymbol{Z}^q\times g$. For simplicity of notation we shall write $G=H\times g$ and denote the elements of G by $x=s\xi$, $y=t\eta$ where $s,t\in H$, $\xi,\eta\in g$. The homomorphism $s\xi\to s$ of $H\times g$ to H we denote by π_H. Being the direct product of groups, the group G has $\mathfrak{H}\times\gamma$ for its generalized character group with γ a discrete group.

Definition 4.5. For each $\sigma\in\gamma$ we define f_σ, μ_σ by

$$f_\sigma(s)=\int_g f(s\xi)\langle\xi,\sigma\rangle d\xi,\ \ \mu_\sigma(s)=\int_g\langle\xi,\sigma\rangle d\mu(s\xi)$$

where $f\in\mathcal{E}(G)$, $\mu\in\mathcal{M}_C(G)$.

It is easy to show that $f_\sigma\in\mathcal{E}(H)$ and $\mu_\sigma\in\mathcal{M}_C(H)$ while $(f*\mu^*)_\sigma=f_\sigma*\mu_\sigma{}^*$. The invariant subspaces of $\mathcal{E}(G)$, $\mathcal{E}(H)$ generated by functions satisfying

$$f*\mu^*=0 \text{ on } G,\quad h*\mu_\sigma{}^*=0 \text{ on } H$$

will be denoted by τ_μ, $\tau_\mu{}^\sigma$ respectively. Thus $f_\sigma\in\tau_\mu{}^\sigma$ if $f\in\tau_\mu$. Any exponential polynomial $P(\cdot,\lambda)$ on G must be of the form $P(x,\lambda)=P(\chi_j(s))(s,\lambda')\langle\xi,\sigma\rangle$, where $\lambda'\in\mathfrak{H}$ and $\sigma\in\gamma$, since the real characters vanish on any compact group.

LEMMA 4.6. *If* $Q(\cdot,\lambda')\in sp(\tau_\mu{}^\sigma)$ $(\lambda'\in\mathfrak{H})$ *then the exponential polynomial* $P(\cdot,\lambda'\sigma)$ *defined by*

$$P(x, \lambda'\sigma) = Q(s, \lambda')\langle \xi, \sigma \rangle$$

belongs to $sp(\tau_\mu)$.

Proof. We have

$$\int_G P(xy, \lambda'\sigma)\, d\mu(y) = \int_H Q(st, \lambda')\{\int_g \langle \xi\eta, \sigma \rangle d\mu(t\eta)\}$$

$$= \langle \xi, \sigma \rangle \int_H Q(st, \lambda')\, d\mu_\sigma(t) = \langle \xi, \sigma \rangle \{Q(\cdot, \lambda') * \mu_\sigma{}^*\}.$$

The assertion of the lemma is now obvious.

The following approximation theorem is, roughly speaking, Fejer's theorem for vector-valued functions.

THEOREM 4.7. *Let* $f \in \mathscr{B}(G)$. *Then, given any* $\epsilon > 0$ *and any compact set* K[3] *in* G, *there exist complex numbers* a_σ, *only finitely many of which are non-zero, such that*

$$\sup_{x \in K} | f(x) - \sum_{\sigma \in \gamma} a_\sigma \langle \xi, \sigma \rangle f_\sigma(s) | < \epsilon.$$

Proof. There is a symmetric neighborhood U of the identity in g for which $\sup_{x \in K} | f(x) - f(x\eta) | < \epsilon$ when $\eta \in U$. Let ψ_1 be any non-negative continuous function with support in U such that $\int_g \psi_1(\eta)\, d\eta = 1$. Then

$$| f(x) - \int_g f(s\eta)\psi_1(\xi\eta^{-1})\, d\eta |$$

$$= | \int_g \{f(x)\psi_1(\xi\eta^{-1}) - f(s\eta)\psi_1(\xi\eta^{-1})\} d\eta |$$

$$\leqq \int_g | f(s\xi) - f(s\eta) | \psi_1(\xi\eta^{-1})\, d\eta < \epsilon, \qquad x \in K.$$

We can find a function $\psi_2 \in L^1(g)$ whose Fourier Transform has compact support on γ and $\int_g | \psi_1(\eta) - \psi_2(\eta) | d\eta < \epsilon$ (see [8] p. 51). We have, therefore,

$$| f(x) - \int_g f(s\eta)\psi_2(\xi\eta^{-1})\, d\eta |$$

$$\leqq | f(x) - \int_g f(s\eta)\psi_1(\xi\eta^{-1})\, d\eta |$$

---

[3] There is no loss in generality in assuming $K = \pi_H{}^{-1}(K')$ for some compact set K' in H.

$$+ \left| \int_g f(s\eta)\{\psi_1(\xi\eta^{-1}) - \psi_2(\xi\eta^{-1})\} d\eta \right|$$
$$< \epsilon(1 + \sup_{y \in xg} |f(y)|), \qquad x \in K.$$

Now ψ_2 must be of the form

$$\psi_2(\xi) = \sum_{\sigma \in \gamma} a_\sigma \langle \xi, \sigma \rangle,$$

where only finitely many a_σ's are non-zero. Thus we have

$$\sup_{x \in K} |f(x) - \sum_{\sigma \in \gamma} a_\sigma \langle \xi, \sigma \rangle f_\sigma(s)| < \epsilon.$$

This completes the proof.

The proof of Theorem 3.2 for compactly generated groups can now be given.

THEOREM 4.8. *Let G be a compactly generated group and let $\mu \in \mathfrak{M}_C(G)$ be given. Then $sp(\tau_\mu)$ is dense in τ_μ.*

Proof. For any $\nu \in \mathfrak{M}_C(G)$ satisfying $P(\cdot, \lambda) * \nu^* = 0$ *on G whenever* $P(\cdot, \lambda) \in sp(\tau_\mu)$, Lemma 4.6 ensures that $Q(\cdot, \lambda') * \nu_\sigma^* = 0$ on H for all $Q(\cdot, \lambda') \in sp(\tau_\mu^\sigma)$. Consequently $h * \nu_\sigma^* = 0$ when $h \in \tau_\mu^\sigma$ (Theorem 4.4). But, for any $f \in \tau_\mu$ and $\sigma \in \gamma$, f_σ belongs to τ_μ^σ and

$$\int_G \langle \eta\xi, \sigma \rangle f_\sigma(st)\, d\nu(t\eta)$$
$$= \int_H f_\sigma(st) \{ \int_g \langle \eta\xi, \sigma \rangle d\nu(t\eta) \} = \langle \xi, \sigma \rangle f_\sigma * \nu_\sigma^* = 0.$$

Hence, applying Theorem 4.7, we deduce

$$|(f * \nu^*)(x)| = |(f * \nu^*)(x) - (\sum_{\sigma \in \gamma} a_\sigma \langle \xi, \sigma \rangle f_\sigma) * \nu^*(x)|$$
$$\leqq \sup_{y \in K} |f(x) - \sum_{\sigma \in \gamma} a_\sigma \langle \xi, \sigma \rangle f_\sigma(s)| \cdot \int_G |d\nu|$$

where K is a compact set depending on x and the support of ν. Thus $f * \nu^* = 0$ on G. This completes the proof of the theorem.

Since we have established Theorem 3.2 for compactly generated groups we can now apply Theorem 3.4 to obtain at once the proof of Theorem 3.2 in general.

5. The proofs given in this paper can be used (with only minor modifications) to yield a result corresponding to Theorem 3.2 for other spaces of functions on groups. For example if $L_{\mathrm{loc}}^P(G)$ denotes the space of functions p-th power integrable ($1 < p < \infty$) on compact sets of G and $L_0{}^q(G)$ denotes its dual then we have a similar result (for definitions see [10] p. 908). The proofs go through virtually unchanged except that a little care is needed when regarding functions integrable on a subgroup H of G as integrable functions on G. This problem (see for example Lemma 4.3) can always be avoided by convolving the function on H with a suitable function on G.

Added in proof. Theorem 3.2 has also been proved independently by R. J. Elliott (see Proc. Camb. Phil. Soc., July (1965)).

University of California,
Berkeley 4, California.
University of Newcastle,
Newcastle-upon-Tyne,
England.

REFERENCES.

[1] J. Braconnier, *L'analyse harmonique dans les groupes abelien*, L'Enseignement Mathématique No. 5, Genève, 1955.

[2] L. Ehrenpreis, "Mean-periodic functions. I," *American Journal of Mathematics*, vol. 77 (1955), pp. 293-328.

[3] M. Lefranc, "Analyse harmonique dans Z^n," *Comptes Rendus de l'Academie des Sciences*, Paris, t. 246 (1958), pp. 1951-1953.

[4] G. Mackey, "Laplace Transform for locally compact groups," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 34 (1948), pp. 156-162.

[5] B. Malgrange, "Existence et approximations des solutions des équations aux derivées partielles et des équations de convolution," *Annales de L'Institut Fourier*, t. 6 (1955-56), pp. 271-355.

[6] ———, "Sur les équations de convolution," *Rendiconti del Seminario Matematico di Torino*, vol. 19 (1959), pp. 19-27.

[7] H. Reiter, "Contributions to Harmonic Analysis, III," *Journal of the London Mathematical Society*, vol. 32 (1957), pp. 477-483.

[8] W. Rudin, *Fourier Analysis on Groups*, New York, Inter-science, 1962.

[9] L. Schwartz, *Théorie des Distributions*, vol. I, Paris, 1950.

[10] ———, "Théorie générale des fonctions Moyenne-Periodique," *Annals of Mathematics*, Series II, vol. 48 (1947), pp. 857-929.

[11] A. Weil, *L'Intégration dans les groupes topologiques*, Actualités Scientifiques et Industrielles, No. 1147, Paris, 1953.

COHOMOLOGICAL PROPERTIES OF COMPLEX SPACES WHICH CARRY OVER TO NORMALIZATIONS.

By Vinicio Villani.

Introduction. To every (reduced) complex space X, with structure sheaf $\mathfrak{O}$, corresponds a *normalization* X^*, which is a normal complex space, with structure sheaf $\mathfrak{O}^*$. There is a natural surjective proper holomorphic map $\pi\colon X^* \to X$, with discrete fibers (see for instance [4]).

This paper is devoted to the study of some properties of the cohomology groups of X, with coefficients in an arbitrary coherent sheaf, which carry over to the normalization X^*.

Let us recall first a theorem known already in the literature:

Theorem 1. *For a given coherent sheaf* $\mathfrak{F}$ *on* X^*, *let* $\pi_0(\mathfrak{F})$ *denote the* 0-*th direct image of* $\mathfrak{F}$ *on* X. *The sheaf* $\pi_0(\mathfrak{F})$ *is coherent, and there is a natural isomorphism*

$$H^j(X^*, \mathfrak{F}) \cong H^j(X, \pi_0(\mathfrak{F})) \qquad \textit{for every } j.$$

Hence, if for some values of j the cohomology groups $H^j(X, \mathfrak{G})$ of X with coefficients in an arbitrary coherent sheaf $\mathfrak{G}$ are zero or are finite dimensional complex vector spaces, we may deduce informations for the cohomology groups of X^* with coefficients in an arbitrary coherent sheaf $\mathfrak{F}$ on X^*.

There are three useful criteria of differential type, stated in [1], on the dimensionality of the cohomology groups of a complex space X with coefficients in an arbitrary coherent sheaf $\mathfrak{G}$; namely:

(I) The following property holds for a *strongly q-pseudoconvex* complex space X:

(A_q) *If* $\mathfrak{G}$ *is an arbitrary coherent sheaf on* X, *then* $\dim_{\mathbf{C}} H^j(X, \mathfrak{G}) < +\infty$, *for every* $j \geqq q$.

(II) The following property holds for a *q-complete* complex space X:

(B_q) *If* $\mathfrak{G}$ *is an arbitrary coherent sheaf on* X, *then* $H^j(X, \mathfrak{G}) = 0$ *for every* $j \geqq q$.

(III) The following property holds for a *strongly q-pseudoconcave* complex space X:

Received June 18, 1965

($\mathrm{C_q}$) *If* $\mathfrak{S}$ *is an arbitrary coherent sheaf on* X, *then* $\dim_{\mathbf{C}} H^j(X, \mathfrak{S}) < +\infty$, *for every* $j < \operatorname*{prof}_X(\mathfrak{S}) - q$.

By definition, $\operatorname*{prof}_X \mathfrak{S} = \inf_{x \in X}(\operatorname{prof}_{\mathfrak{O}_x} \mathfrak{S}_x)$. In § 2 we recall the definition of $\operatorname{prof}_{\mathfrak{O}_x} \mathfrak{S}_x$, and more generally of $\operatorname{prof}_A F$, F being an A-module of finite type over a local noetherian ring A.[1]

For the definitions of strongly q-pseudoconvex, q-complete and strongly q-pseudoconcave spaces, see § 4.

Let us turn back to the cohomology groups of X and X^*. An immediate consequence of Theorem 1 is

COROLLARY 1. *If for a certain integer* q *the property* ($\mathrm{A_q}$), *or the property* ($\mathrm{B_q}$), *holds for* X, *the same property holds also for* X^*.

A similar result holds good also for the property ($\mathrm{C_q}$) (see Corollary 2, below); but in order to show this, one needs the following algebraic theorem:

THEOREM 2. *Let* A, B *be two local noetherian rings,* B *being an integral extension of finite type of* A. *Let* F *be a* B*-module of finite type. The natural injection map* $A \rightarrow B$ *induces on* F *a structure of an* A*-module of finite type, and one has:*

$$\operatorname{prof}_B F = \operatorname{prof}_A F.$$

Applying this theorem to the stalks of the sheaves $\mathfrak{F}$ on X^* and $\pi_0(\mathfrak{F})$ on X, we obtain, for every $x \in X$:

$$\operatorname{prof}_{\mathfrak{O}_x}(\pi_0(\mathfrak{F}))_x = \min_{\{x_i \in X^*;\, \pi(x_i) = x\}} (\operatorname{prof}_{\mathfrak{O}^*_{x_i}} \mathfrak{F}_{x_i}),$$

whence: $\operatorname*{prof}_X \pi_0(\mathfrak{F}) = \operatorname*{prof}_{X^*} \mathfrak{F}$. This last result, together with Theorem 1, proves the

COROLLARY 2. *If for a certain integer* q *the property* ($\mathrm{C_q}$) *holds for* X, *the same property holds also for* X^*.

Results, similar to those proved in Corollaries 1 and 2, could be obtained in a more geometric way, if we prove the following: If X is strongly q-pseudoconvex (or q-complete, or strongly q-pseudoconcave), the same is true also for X^*. But in this connection I succeeded in proving only the following

THEOREM 3. *Suppose the* (*analytic*) *set* N *of non-normal points of* X *has dimension zero. Then, if* X *is strongly* q*-pseudoconvex* (*or* q*-complete, or strongly* q*-pseudoconcave*), *the same is true also for* X^*.

---

[1] Notice that in [1] the symbol $\operatorname{dih}(\mathfrak{S})$ is used instead of $\operatorname*{prof}_X \mathfrak{S}$. In § 2 we shall show that, in fact, $\operatorname*{prof}_X \mathfrak{S}$ and $\operatorname{dih}(\mathfrak{S})$ are the same.

The problem of extending this result to the general case, that is to the case when N has arbitrary dimension, remains unsolved.

The author is very much indebted to Professor Zariski for his valuable suggestions concerning the proof of Theorem 2.

1. Let $\pi\colon Y \to X$ be a holomorphic map of a complex space Y into a complex space X. Let $\mathfrak{F}$ be an analytic sheaf on Y, and ν a non-negative integer. Then, by definition, the *ν-th direct image* of $\mathfrak{F}$ is the analytic sheaf $\pi_\nu(\mathfrak{F})$ on X, associated to the presheaf $\{T_\nu(U), r_V{}^U\}$, where $T_\nu(U) = H^\nu(\pi^{-1}(U), \mathfrak{F})$ and $r_V{}^U$ are the natural restriction homomorphisms.

The following theorem is proved in [4] (Th. 27):

If π is a proper map with discrete fibers, then we have for an arbitrary coherent sheaf $\mathfrak{F}$ on Y:

$$\pi_\nu(\mathfrak{F}) = 0 \textit{ if } \nu \geqq 1;$$
$$\pi_0(\mathfrak{F}) \textit{ is a coherent sheaf on } X.$$

On the other hand:

If $\pi_\nu(\mathfrak{F}) = 0$ for every $\nu \geqq 1$, then we have an isomorphism:

$$H^j(Y, \mathfrak{F}) \cong H^j(X, \pi_0(\mathfrak{F})) \textit{ for every } j.$$

(See [3], Th. 6).[2]

These results imply Theorem 1, by taking for Y the normalization X^* of X.

2. Let A be a local (noetherian) ring with maximal ideal $\mathfrak{m}_A$ (or briefly $\mathfrak{m}$) and F be an A-module of finite type. A sequence $x_1, \cdots, x_r$ of elements of $\mathfrak{m}$ is called a *regular F-sequence* (of length r) if x_i is not a zero divisor with respect to $F/(x_1, \cdots, x_{i-1})F$, $(i = 1, \cdots, r)$.

Define

$$F_{i-1} = (x_1, \cdots, x_{i-1})F$$

and

$$F_{i-1} : (x_i) = \{f \in F; x_i f \in F_{i-1}\}.$$

Clearly $x_1, \cdots, x_r$ is a regular F-sequence if and only if

---

[2] One can give a direct proof of this last statement: consider the spectral sequence associated to the continuous map $\pi\colon Y \to X$ (see [2], p. 201); $E_2{}^{p,q}$ is isomorphic to $H^p(X, \pi_q(\mathfrak{F}))$. Since $\pi_q(\mathfrak{F}) = 0$ for every $q \geqq 1$, it follows that $E_2{}^{p,q} = 0$ if $q \neq 0$, so that the spectral sequence degenerates and we get isomorphisms

$$H^p(Y, \mathfrak{F}) \cong H^p(X, \pi_0(\mathfrak{F}))$$

for every p.

$$F_{i-1}:(x_i)=F_{i-1} \qquad (i=1,\cdots,r).$$

We set: $\mathrm{prof}_A F=$ *the largest integer* r, *such that there exists a regular* F*-sequence of length* r (remark that $\mathrm{prof}_A F=+\infty$ if and only if $F=0$).

Note. The symbol $\mathrm{prof}_A F$ has been introduced only recently in the literature; the same number is called $\mathrm{codh}_A F$ in [7].

Now we recall the definition of $\mathrm{dih}(F)$, as introduced in [1] for modules F over a local ring A which is a quotient of a regular local ring R: the natural projection $\phi: R\to A$ induces an R-module structure on F. We suppose $F\neq 0$; let d be the least length of a free resolution of F as an R-module:

$$0\to R^{p_d}\to R^{p_{d-1}}\to\cdots\to R^{p_0}\to F\to 0,$$

and let m be the Krull dimension of R. Then, by definition, $\mathrm{dih}(F)=m-d$. We shall now show that for an F as above $\mathrm{prof}_A F$ coincides with $\mathrm{dih}(F)$. We observe that $m-d=\mathrm{prof}_R F$ ([7], chap. IV, p. 31). Now, since ϕ: $R\to A$ is a surjective homomorphism of local rings, $\mathrm{prof}_R F=\mathrm{prof}_A F$ ([7], chap. IV, p. 16), and this concludes the proof.

There is another characterization for $\mathrm{prof}_A F$ (see [7], chap. IV, pp. 12-13): set $k=A/\mathfrak{m}$; then $\mathrm{prof}_A F$ is the least integer n, such that $\mathrm{Ext}^n{}_A(k,F)\neq 0$.

LEMMA 1. *Let* $F=F_1\oplus\cdots\oplus F_s$ *be a direct sum of* A*-modules of finite type. Then* $\mathrm{prof}_A F=\min\limits_{j=1,\cdots,s}(\mathrm{prof}_A F_j)$.

The proof of Lemma 1 follows from

$$\mathrm{Ext}^n{}_A(k,F_1\oplus\cdots\oplus F_s)=\mathrm{Ext}^n{}_A(k,F_1)\oplus\cdots\oplus\mathrm{Ext}^n{}_A(k,F_s).$$

The proof of Theorem 2 makes use of the following two lemmas:

LEMMA 2. *Let* F *be a* B*-module,* G *a submodule of* F, *and* x *a fixed element in* $\mathfrak{m}_B$. *Then* $G:(x)=G$ *implies* $G:(x^k)=G$ *for every integer* $k\geqq 1$.

Proof. $G:(x^k)=G:(x^{k-1})=\cdots=G:(x)=G$.

LEMMA 3. *Let* F *be a* B*-module of finite type and* G *a submodule of* F. *Let* $G=\bigcap\limits_{j=1}^{h}M_j$ *be an irredundant decomposition of* G *by primary* B*-modules* M_j (see [5], p. 24, of [7], chap. I, p. 14). *Let* $\mathfrak{q}_j=\{b\in B; bF\subseteq M_j\}$ *be the primary ideals of* B *associated to the modules* M_j *and let* $\mathfrak{p}_j$ *be the corresponding prime ideals of* B.

For a fixed element $z\in\mathfrak{m}_B$ *one has* $G:(z)=G$ *if and only if* $z\notin\bigcup\limits_{j=1}^{h}\mathfrak{p}_j$.

Proof. (a) Assume $G=G:(z)$. We prove that $z\notin\mathfrak{p}_1$ (similarly one can show that $z\notin\mathfrak{p}_2,\cdots,z\notin\mathfrak{p}_h$). In fact let, if possible, $z\in\mathfrak{p}_1$. Then there exists an integer k, such that $z^k\in\mathfrak{q}_1$, whence $\mathfrak{q}_1F:(z^k)=F$. The obvious inclusion $M_1\supseteq\mathfrak{q}_1F$ implies

$$M_1:(z^k)\supseteq\mathfrak{q}_1F:(z^k)=F,$$

and therefore $M_1:(z^k)=F$. Hence

$$G:(z^k)=\bigcap_{j=1}^{h}(M_j:(z^k))=\bigcap_{j=2}^{h}(M_j:(z^k))\,;$$

on the other hand, Lemma 2 implies $G:(z^k)=G=\bigcap_{j=1}^{h}M_j$; therefore $\bigcap_{j=2}^{h}(M_j:(z^k))=\bigcap_{j=1}^{h}M_j$. But since $M_j\subseteq M_j:(z^k)$, the inclusion

$$\bigcap_{j=2}^{h}M_j\subseteq\bigcap_{j=2}^{h}(M_j:\ (z^k))=\bigcap_{j=1}^{h}M_j$$

follows. Since the opposite inclusion $\bigcap_{j=2}^{h}M_j\supseteq\bigcap_{j=1}^{h}M_j$ is obvious, we have finally

$$G=\bigcap_{j=1}^{h}M_j=\bigcap_{j=2}^{h}M_j,$$

contradicting the irredundancy of the decomposition $G=\bigcap_{j=1}^{h}M_j$. (b) Assume $z\notin\bigcup_{j=1}^{h}\mathfrak{p}_j$. We have then to show that $G:(z)=G$. Equivalently we may show that $M_j:(z)=M_j$ for $j=1,\cdots,h$. And in fact, suppose this equality doesn't hold for some j (for instance $j=1$); then z is a zero divisor with respect to F/M_1. But since M_1 is a primary module, this implies that z is nilpotent with respect to F/M_1, i.e. there exists an integer k such that $z^kF\subseteq M_1$. But in this case $z^k\in\mathfrak{q}_1$, whence $z\in\mathfrak{p}_1$, contradicting our assumption.

Now we are in a position to prove Theorem 2. We first show that $\mathrm{prof}_B\,F\leqq\mathrm{prof}_A\,F$.

Let $x_1,\cdots,x_r$ be a regular sequence of F, as a B-module, and let $F_{r-1}=\bigcap_{j=1}^{h}M_j$ be an irredundant decomposition of the B-module

$$F_{r-1}=(x_1,\cdots,x_{r-1})F.$$

If $\mathfrak{p}_j$ is the prime ideal of B, corresponding to the primary module M_j, one has $\bigcup_{j=1}^{h}\mathfrak{p}_j\subsetneq\mathfrak{m}_B$; in fact, from $F_{r-1}:(x_r)=F_{r-1}$ we deduce that $x_r\notin\bigcup_{j=1}^{h}\mathfrak{p}_j$, according to Lemma 3, while $x_r\in\mathfrak{m}_B$ by hypothesis.

Define $\mathfrak{p}'_j=\mathfrak{p}_j\cap A$; these are prime ideals of A. The "Going-up theorem" (see [8], p. 259) implies $\bigcup_{j=1}^{h} \mathfrak{p}'_j \subsetneq \mathfrak{m}_A$. Hence there exists an element $z \in \mathfrak{m}_A$, such that $z \notin \bigcup_{j=1}^{h} \mathfrak{p}'_j$. Consider z as an element of $\mathfrak{m}_B$; again Lemma 3 implies that $x_1, \cdots, x_{r-1}, z$ is a new regular sequence of F as a B-module, and in this sequence the last element belongs to $\mathfrak{m}_A$. Since a permutation on the elements of a regular F-sequence is again a regular F-sequence ([5], p. 96), it follows that $z, x_1, \cdots, x_{r-1}$ is a regular sequence of F, as a B-module, and in this sequence the first element belongs to $\mathfrak{m}_A$. Repeating the same process r times we conclude that *if there exists a regular sequence* $x_1, \cdots, x_r$ *of* F, *considered as a* B-*module* (*i.e. with* $x_j \in \mathfrak{m}_B$), *then there exists also a regular sequence* $z_1, \cdots, z_r$ *of* F, *still considered as a* B-*module, with* $z_j \in \mathfrak{m}_A$ $(j = 1, \cdots, r)$. But then $z_1, \cdots, z_r$ is a regular sequence of F also as an A-module. The inequality $\operatorname{prof}_B F \leqq \operatorname{prof}_A F$ follows from this.

It remains now to prove the opposite inequality: $\operatorname{prof}_B F \geqq \operatorname{prof}_A F$. Let $x_1, \cdots, x_r$ be a regular sequence of F, as an A-module (i.e. with $x_j \in \mathfrak{m}_A$); it suffices to show that $x_1, \cdots, x_r$, considered as elements of $\mathfrak{m}_B$, still form a regular sequence of F, as a B-module. And in fact, F_{i-1} as well as $F_{i-1} : (x_i)$ do not change if we substitute the ring of coefficients A with the ring B; therefore if the condition $F_{i-1} = F_{i-1} : (x_i)$ is verified for all i, with respect to the A-module structure of F, the same condition is verified also with respect to the B-module structure of F.

3. From the theorem now proved, we deduce the Corollary 2, stated in the Introduction.

For a given $x \in X$, let x_i $(i = 1, \cdots, s)$ denote the points of X^*, such that $\pi(x_i) = x$. Let $\mathfrak{F}_{x_i}$ denote the stalk over x_i of the coherent sheaf $\mathfrak{F}$; each $\mathfrak{F}_{x_i}$ is an $\mathfrak{O}^*_{x_i}$-module of finite type. The stalk $(\pi_0(\mathfrak{F}))_x$ over x of the sheaf $\pi_0(\mathfrak{F})$ is an $\mathfrak{O}_x$-module of finite type.

LEMMA 4. *With these notations we have*:

$$\operatorname{prof}_{\mathfrak{O}_x}(\pi_0(\mathfrak{F}))_x = \min_{i=1,\cdots,s} (\operatorname{prof}_{\mathfrak{O}^*_{x_i}} \mathfrak{F}_{x_i}).$$

Proof. There is a decomposition of the germ of X at x, as a union of s germs of irreducible analytic sets $A_1, \cdots, A_s$, with structure sheaves $\mathfrak{O}^{(1)}, \cdots, \mathfrak{O}^{(s)}$. After renumbering the indices, the normalization of each A_i is (isomorphic to) the germ of X^* at x_i. The stalk $\mathfrak{O}^*_{x_i}$ is the integral closure of the corresponding stalk $\mathfrak{O}^{(i)}_x$. Hence each $\mathfrak{O}^*_{x_i}$-module $\mathfrak{F}_{x_i}$ may be endowed also with an $\mathfrak{O}^{(i)}_x$-module structure, and according to Theorem

2 one has $\operatorname{prof}_{\mathfrak{O}^*_{x_i}} \mathfrak{F}_{x_i} = \operatorname{prof}_{\mathfrak{O}^{(i)}_x} \mathfrak{F}_{x_i}$. Furthermore, there are natural surjective homomorphisms $\phi_i \colon \mathfrak{O}_x \to \mathfrak{O}^{(i)}_x$, endowing every $\mathfrak{F}_{x_i}$ with an $\mathfrak{O}_x$-module structure, and by a result already mentioned above ([7], chap. IV, p. 16) we get $\operatorname{prof}_{\mathfrak{O}^{(i)}_x} \mathfrak{F}_{x_i} = \operatorname{prof}_{\mathfrak{O}_x} \mathfrak{F}_{x_i}$. In order to accomplish the proof of Lemma 4, it suffices to notice that one has a natural isomorphism of $\mathfrak{O}_x$-modules: $(\pi_0(\mathfrak{F}))_x \cong \mathfrak{F}_{x_1} \oplus \cdots \oplus \mathfrak{F}_{x_s}$, whence, applying Lemma 1, we obtain:

$$\operatorname{prof}_{\mathfrak{O}_x}(\pi_0(\mathfrak{F}))_x = \min_{i=1,\cdots,s} (\operatorname{prof}_{\mathfrak{O}_x} \mathfrak{F}_{x_i}) = \min_{i=1,\cdots,s} (\operatorname{prof}_{\mathfrak{O}^*_{x_i}} \mathfrak{F}_{x_i}).$$

In order to prove the Corollary 2, we remark that from the definition of $\operatorname*{prof}_X \mathfrak{G}$, and from Lemma 4 we get

$$\operatorname*{prof}_X \pi_0(\mathfrak{F}) = \inf_{x \in X} \operatorname{prof}_{\mathfrak{O}_x}(\pi_0(\mathfrak{F}))_x = \inf_{x \in X} \Big(\min_{i=1,\cdots,s} (\operatorname{prof}_{\mathfrak{O}^*_{x_i}} \mathfrak{F}_{x_i}) \Big)$$

$$= \inf_{x^* \in X^*} (\operatorname{prof}_{\mathfrak{O}^*_{x^*}} \mathfrak{F}_{x^*}) = \operatorname*{prof}_{X^*} \mathfrak{F}.$$

Corollary 2 is now an obvious consequence of Theorem 1.

4. We recall the definitions of a strongly q-pseudoconvex, q-complete, strongly q-pseudoconcave complex space.

Let ϕ be a continuous, real valued function on a complex space X. For a fixed $x \in X$, let W_x denote a neighborhood of x, isomorphic to an analytic set A, defined on an open set U of some $\boldsymbol{C}^N$ (with complex coordinates $z_1, \cdots, z_N$); let ϕ_0 denote the image of ϕ on the analytic set A; ϕ_0 is said to be differentiable of class C^∞ on A, if there exists on U a function $\hat{\phi}$, differentiable of class C^∞, such that $\hat{\phi} \mid A = \phi_0$. The function ϕ_0 is said to be (weakly) q-plurisubharmonic[3] on A, if $\hat{\phi}$ can be chosen so that the hermitian form in N variables:

$$\mathscr{L}(\hat{\phi}, p) = \sum_{\alpha,\beta=1}^{N} \left(\frac{\partial^2 \hat{\phi}}{\partial z_\alpha \partial \bar{z}_\beta} \right)_p u_\alpha \bar{u}_\beta$$

has at least $N - q + 1$ eigenvalues $\geqq 0$ at every point $p \in U$. The function ϕ_0 is said to be strongly q-plurisubharmonic on A, if $\hat{\phi}$ can be chosen such that all these eigenvalues are > 0 at every point $p \in U$.

The function ϕ on the complex space X is said to be differentiable of class C^∞, or q-plurisubharmonic, or strongly q-plurisubharmonic, if for all $x \in X$ the corresponding functions ϕ_0 on the analytic sets which give local realizations of X as a subset of some $\mathbf{C}^N$ (N possibly depending on the point x) have the same properties.

---

[3] More recently the function ϕ_0 is usually called q-plurisubharmonic if the corresponding form $\mathscr{L}(\hat{\phi}, p)$ has at least $N - q$ eigenvalues $\geqq 0$; accordingly, the definitions of q-pseudoconvexity, q-completeness and q-pseudoconcavity are shifted by one.

The complex space X is said to be *strongly q-pseudoconvex* if there exists a compact set K in X, and a real valued function ϕ on X, which is differentiable of class C^∞ on X, strongly q-plurisubharmonic on $X-K$, and such that for every $c \in \mathbf{R}$ the sets $B_c = \{x \in X; \phi(x) < c\}$ are relatively compact in X.

If in the above definition one can choose $K = \emptyset$, the complex space X is said to be *q-complete.*

The complex space X is said to be *strongly q-pseudoconcave* if there exists a compact set K in X, and a non-negative real valued function ϕ on X, which is differentiable of class C^∞ on X, strongly q-plurisubharmonic on $X-K$, and such that for every $c > 0$ the sets $B_c = \{x \in X; \phi(x) > c\}$ are relatively compact in X.

The proof of Theorem 3 makes use of

LEMMA 5. *Let Z be a complex space, H a compact subset of Z and $\{p_s\}$ $(s = 1, 2, \cdots)$ a countable set of isolated points in $Z-H$. Assume there exists on Z a real valued function ψ, which is differentiable of class C^∞ on Z, strongly q-plurisubharmonic on $Z-(H \cup \{p_s\})$ and weakly q-plurisubharmonic at the points p_s. Assume furthermore that ψ satisfies one of the following conditions*: (1) *The sets $B_c = \{z \in Z; \psi(z) < c\}$ are relatively compact in Z for every real c, or* (2) *ψ is non-negative and the sets $B_c = \{z \in Z; \psi(z) > c\}$ are relatively compact in Z for every real positive c.*

Then there exists on Z a real valued function $\lambda \geqq \psi$, which is differentiable of class C^∞ on Z, strongly q-plurisubharmonic on $Z-H$, and such that in the case (1) *the sets $S_c = \{z \in Z; \lambda(z) < c\}$ are relatively compact in Z for every real c, in the case* (2) *the sets $S_c = \{z \in Z; \lambda(z) > c\}$ are relatively compact in Z for every real positive c.*

Hence, with respect to the function λ, the space Z is strongly q-pseudoconvex (q-complete if $H = \emptyset$) in the case (1), *and strongly q-pseudoconcave in the case* (2).

Proof. We consider first the case (2). Let $\{U_s\}$ be a collection of neighborhoods of the $\{p_s\}$, such that

(I) for every s, U_s is a relatively compact subset of $Z-H$;

(II) if $s \neq t$, $\bar{U}_s \cap \bar{U}_t = \emptyset$;

(III) for every s, $\bar{U}_s$ is contained in an open set $W_s \subset Z$, which is isomorphic to an analytic set A_s defined on an open set of some $\mathbf{C}^{N_s}$;

(IV) for every s, $\psi(z) < 2\psi(p_s)$ at every point $z \in \bar{U}_s$.

Then, for every s, there exists on Z a non-negative real valued function χ_s, which is differentiable of class C^∞ on Z, strongly 1-plurisubharmonic on a suitable neighborhood V_s of p_s, and zero on $Z-U_s$: it suffices to consider on A_s the function $\sum_{\alpha=1}^{N_s} | z_\alpha |^2$, to take its inverse image on W_s, and to smooth down differentiably to zero this function outside a convenient neighborhood of p_s.

Moreover, for every s there exists a suitable small positive constant k_s, such that

(i) the (differentiable) function $\psi + k_s\chi_s$ is strongly q-plurisubharmonic on $\bar{U}_s - V_s$;

(ii) $\sup_{z \in Z} k_s\chi_s(z) < \psi(p_s)$.

By construction, the differentiable function

$$\lambda = \psi + \sum_s k_s\chi_s$$

is non-negative, and strongly q-plurisubharmonic at every point of $Z-H$. It remains only to show that the sets $S_c = \{z \in Z; \lambda(z) > c\}$ are relatively compact in Z, for every $c > 0$. Actually, it suffices to show that this assertion is true, when c describes a sequence of positive numbers $\{c_r\}$, converging towards zero; for instance $c_r = 1/r$ $(r = 1, 2, \cdots)$. Let $p_1, \cdots, p_t$ be the points of the set $\{p_s\}$, belonging to $B_{c_{3r}}$. The closed set $\bar{S}_{c_r}$ is contained in the compact set $B_{c_r} \cup (\bar{U}_1 \cup \cdots \cup \bar{U}_t)$, as follows from the hypothesis (IV) on U_s, and from the hypothesis (ii) on the constants k_s. Hence S_{c_r} is relatively compact in Z. This achieves the proof of Lemma 5 in the case (2).

The same arguments apply also for the case (1); one should only remark that now the hypotheses (IV) on U_s and (ii) on k_s are superfluous, since $\lambda(z) < c$ implies $\psi(z) < c$, whence it follows that the closed set $\bar{S}_c$ is contained in the compact set $\bar{B}_c$, for every c.

Proof of Theorem 3. By definition there exists a function ϕ on X, ensuring the strong q-pseudoconvexity (or q-completeness, or strong q-pseudoconcavity) of X. Consider on X^* the function $\psi =$ inverse image of ϕ by the map $\pi: X^* \to X$, and set $H = \pi^{-1}(K)$. At every point $z \in X^* - H$, ψ is weakly q-plurisubharmonic (the proof is a straightforward calculation; one has to notice only that if $\pi: A^* \to A$ is a holomorphic map of a germ of analytic set A^* defined in a $\mathbf{C}^M$ into a germ of analytic set A defined in a $\mathbf{C}^N$, then π may be extended to a holomorphic map of some open set of $\mathbf{C}^M$ containing the germ A^*, into an open set of $\mathbf{C}^N$ containing the germ A).

Furthermore if $z \in X^* - H$, and $z \notin \pi^{-1}(N)$, the function ψ is also strongly q-plurisubharmonic at z; but if $z \in \pi^{-1}(N)$, the function ψ may fail to be strongly q-plurisubharmonic at z. Since π is a proper map, the condition that for a given c the set $\{x \in X; \phi(x) < c\}$ is relatively compact in X (or the condition that the set $\{x \in X; \phi(x) > c\}$ is relatively compact in X) implies that for the same value of c the set $\{z \in X^*; \psi(z) < c\}$ (or $\{z \in X^*; \psi(z) > c\}$) is relatively compact in X^*. As a consequence of this fact, X^* satisfies the second axiom of countability. Hence $\pi^{-1}(N)$, being a zero-dimensional analytic set of X^*, is a countable set of isolated points; let $\{p_s\}$ $(s = 1, 2, \cdots)$ be those points of $\pi^{-1}(N)$ which do not belong to H.

With these notations we can apply Lemma 5 to the space $Z = X^*$; Theorem 3 follows from this.

SCUOLA NORMALE SUPERIORE, PISA (ITALY).

REFERENCES.

[1] A. Andreotti and H. Grauert, "Théorèmes de finitude pour la cohomologie des espaces complexes," *Bulletin de la Société Mathématique de France,* vol. 90 (1962), pp. 193-259.

[2] R. Godemont, *Topologie algébrique et théorie des faisceaux,* Hermann, 1958.

[3] H. Grauert and R. Remmert, "Bilder und Urbilder analytischer Garben," *Annals of Mathematics,* vol. 68 (1958), pp. 393-443.

[4] ———; "Komplexe Räume," *Mathematische Annalen,* vol. 136 (1958), pp. 245-318.

[5] M. Nagata, *Local Rings,* Interscience Publishers, 1962.

[6] R. Narasimhan, "A note on Stein spaces and their normalizations," *Annali della Scuola Normale Superiore di Pisa,* vol. 16 (1962), pp. 327-333.

[7] J. P. Serre, *Algèbre locale-Multiplicités* (Cours au Collège de France 1957-58).

[8] O. Zariski and P. Samuel, *Commutative Algebra,* vol. I, Van Nostrand, 1958.

ALGEBRAIC INVARIANTS OF ISOTOPY OF LINKS.

By R. Holmes and N. Smythe.

---

1. Introduction. Consider a link L in the interior of an orientable 3-manifold M. It has been shown by J. Milnor [2] that the quotient groups G/G_q, $q = 1, 2, \cdots$, of the lower central series of $G = \pi_1(M - L)$ are invariant under isotopy of L. From these invariants he extracted numerical invariants, the $\bar{\mu}$-invariants.

In this paper, we derive some analogous invariants connected with the integer group ring of G. R. H. Fox noticed that a simplification of the proof of invariance of G/G_q can be obtained if the notion of isotopy of L is replaced by another concept, which we here call F-isotopy (see definition, Section 2). It is shown in [3] that links which are isotopic under a polyhedral isotopy are F-isotopic; thus, in this sense, the invariants of F-isotopy to be described are also invariants of isotopy of links. Throughout, all links are assumed polyhedral.

We show that the rings $JG/\mathfrak{g}^q$, $q = 1, 2, \cdots$, are invariants of L under F-isotopy, where JG is the integer group ring of G, and $\mathfrak{g}$ is the fundamental ideal (the kernel of the trivializer map $JG \to J$).

In the particular case of links of two components in the 3-sphere, we derive numerical invariants of F-isotopy from the Magnus expansion of a longitude (the co-efficients of this expansion define the $\bar{\mu}$-invariants). An application shows that these invariants are stronger than the $\bar{\mu}$-invariants.

2. Let $L = l_1 \cup l_2 \cup \cdots \cup l_n$, $L' = l_1' \cup l_2' \cup \cdots \cup l'_n$ be oriented links in an orientable 3-manifold M.

Definition. L' arises from L by a simple F-isotopy f on the i-th component if there is an orientation preserving homeomorphism $f\colon M \to M$ and a solid torus $V \subset M - (L - l_i)$ such that $f(l_j) = l_j'$ for $j \neq i$, and either

i) l_i is a core of V, and l_i' has winding number 1 in $f(V)$, or

ii) l_i' is a core of $f(V)$, and l_i has winding number 1 in V.

Then two links are said to be F-istotopic in M if one can be obtained from the other by a finite sequence of simple F-isotopies.

Received June 4, 1965.

The purpose of the homeomorphism f is to ensure that links of the same oriented link type are F-isotopic.

Suppose l_i is nullhomologous in M, and let N_i denote a solid torus regular neighbourhood of l_i. We define an i-th longitude-meridian pair (u_i, x_i) for L to be a pair of elements of $G = \pi_1(M - L)$ such that u_i is represented by a simple closed curve on $\mathrm{Bd}\, N_i$ with winding number 1 in N_i and which is nullhomologous in $M - l_i$, together with an arc A joining $\mathrm{Bd}\, N_i$ to the base point of $M - L$, and x_i is represented by a meridian on $\mathrm{Bd}\, N_i$ together with A. These elements are well-determined in G by the arc A and by the orientations of l_i and N_i. Any other i-th longitude-meridian pair (u_i', x_i') for L arises from (u_i, x_i) by simultaneous conjugation, i.e., $u_i' = a u_i a^{-1}$ and $x_i' = a x_i a^{-1}$ for some $a \in G$.

Now suppose L' arises from L by a simple F-isotopy f on the i-th component, in fact we shall suppose that there is a solid torus $V \subset M - (L - l_i)$ with core l_i such that l_i' has winding number 1 in $f(V)$, and $f(l_j) = l_j'$ for $j \neq i$.

Then the following lemma exhibits a relationship between the fundamental groups $G = \pi_1(M - L)$ and $H = \pi_1(M - L')$.

LEMMA. *G is a retract of H under a map $\psi\colon H \to G$ inducing an isomorphism $\psi_*\colon H_1(M - L') \approx H_1(M - L)$.*

Moreover, if l_j is nullhomologous in M, then l_j' is nullhomologous in M, and ψ preserves longitude-meridian pairs (i.e., if (u_j', x_j') is a j-th longitude-meridian pair for L', then $(\psi(u_j'), \psi(x_j'))$ is a j-th longitude-meridian pair for L).

Proof. Let $M_i = M - (L - l_i)$ and $M_i' = M - (L' - l_i')$. By the van Kampen theorem, H is the direct limit of the system

$$\begin{tikzcd} \pi_1(M_i' - f(V)) & \pi_1(f(\mathrm{Bd}\, V)) \arrow[l, "i"'] \arrow[r, "j"] & \pi_1(f(V) - l_i') \end{tikzcd}$$

the maps i, j being induced by inclusion. Now $M_i' - f(V) \cong M_i - V \cong M - L$ since l_i is a core of V. Therefore $\pi_1(M_i' - f(V)) \approx G$. Let $\phi\colon G \to H$, $\alpha\colon \pi_1(f(V) - l_i') \to H$ be the maps defined by the direct limit.

Consider the commutative diagram

$$\begin{tikzcd} \pi_1(f(\mathrm{Bd}\, V)) \arrow[r, "h_1"] \arrow[d, "j"'] & H_1(f(\mathrm{Bd}\, V)) \arrow[d, "j_*"'] \\ \pi_1(f(V) - l_i') \arrow[r, "h_2"] & H_1(f(V) - l_i') \end{tikzcd}$$

The abelianizer h_1 is an isomorphism. $H_1(f(V)-l_i')$ is free abelian generated by a longitude of $f(\mathrm{Bd}\,V)$ and a meridian of l_i'; since l_i' has winding number 1 in $f(V)$, a meridian of l_i' is homologous to a meridian cycle on $f(\mathrm{Bd}\,V)$. Thus j_* is also an isomorphism.

Then we may construct a map of the above system of groups into G, according to the commutative diagram

$$\begin{tikzcd}
 & \pi_1(f(\mathrm{Bd}\,V)) \arrow[dl, "i"'] \arrow[dr, "j"] & \\
G \arrow[dr, "\text{identity}"'] & & \pi_1(f(V)-l_i') \arrow[dl, "i\cdot h_1^{-1}\cdot j_*^{-1}\cdot h_2"] \\
 & G &
\end{tikzcd}$$

By the universal property of direct limits, there is a map $\psi\colon H\rightarrow G$ such that $\psi\phi$ is the identity on G. Thus G is a retract of H.

H is generated by the images of α and ϕ, so $H_1(M-L')$ is generated by the images of the induced maps α_* and ϕ_*. But $\alpha_*=i_*\cdot(h_1)_*^{-1}\cdot j_*^{-1}\cdot\phi_*$, so ϕ_* is onto. On the other hand, $\psi_*\phi_*=1$, so ϕ_* is $1:1$. Thus ϕ_* and ψ_* are isomorphisms of the first homology groups.

Since l_i' is homotopic to $f(l_i)$ in the complement of $L'-l_i'$, l_i' is nullhomologous in M if and only if l_i is nullhomologous in M.

Clearly ψ preserves meridians, and also longitudes if they are defined. (Note that ϕ need not preserve meridians however: the image $\phi(x_i)$ of an i-th meridian is represented by a meridian of $f(\mathrm{Bd}\,V)$, which need not be a meridian of l_i'.)

3. THEOREM. *Let $G=\pi_1(M-L)$, $H=\pi_1(M-L')$ be the groups of F-isotopic links L, L' in the 3-manifold M. Let G_q, H_q denote the q-th groups of the lower central series $(G_1=G, G_q=[G,G_{q-1}])$. Let JG, JH denote the integer group rings of G, H, with fundamental ideals $\mathfrak{G}$, $\mathfrak{H}$. Then there exist group isomorphisms*

$$\phi_q\colon G/G_q\approx H/H_q,\qquad \phi_{pq}\colon G_p/G_q\approx H_p/H_q$$

and ring isomorphisms

$$\phi_q^*\colon JG/\mathfrak{G}^q\approx JH/\mathfrak{H}^q,\qquad \phi_{pq}^*\colon \mathfrak{G}^p/\mathfrak{G}^q\approx\mathfrak{H}^p/\mathfrak{H}^q$$

In the case that the components of L are nullhomologous in the oriented manifold M, then the isomorphisms ϕ_q and ϕ_q^* preserve longitude-meridian pairs.

Proof. By symmetry, we need only consider the case that L' arises from L by a simple F-isotopy on the i-th component, as in the lemma. We use the same notation; thus ϕ and ψ are the maps of the lemma. These maps induce ring homomorphisms ϕ^* and ψ^*, which in turn induce maps $\phi_q{}^*$ and $\psi_q{}^*$ of the quotient rings. We give only the proof that $\phi_q{}^*$ is an isomorphism. The proof for ϕ_q is very similar. That ϕ_{pq} and $\phi_{pq}{}^*$ are isomorphisms will then be immediate corollaries.

$\psi_q{}^*\phi_q{}^*$ is the identity map, so $\phi_q{}^*$ is $1:1$. There is a natural equivalence between $JG/\mathfrak{G}^2$ and $J \oplus G/G_2$ under which $\phi_2{}^*$ corresponds to the map induced by ϕ_2. By the lemma, ϕ_2 is an isomorphism, so $\phi_2{}^*$ is a ring isomorphism.

Now suppose $\phi_p{}^*$ is onto, for each integer $p \leqq q-1$. Then for any $h \in JH$, there is an element $g \in JG$ such that $h = \phi^*(g) + h'$, with $h' \in \mathfrak{H}^{q-1}$. Then $h' = \sum m_j a_j b_j$, where $m_j \in J$, $a_j \in \mathfrak{H}$, $b_j \in \mathfrak{H}^{q-2}$. By assumption, $a_j = \phi^*(g_j) + a_j'$, $b_j = \phi^*(g_j') + b_j'$, where g_j and g_j' are elements of JG, and a_j' and b_j' are in $\mathfrak{H}^{q-1}$.

$$\begin{aligned} \text{Then } h' &= \textstyle\sum m_j(\phi^*(g_j) + a_j')(\phi^*(g_j') + b_j') \\ &= \textstyle\sum m_j\phi^*(g_j)\phi^*(g_j') + h'', \text{ with } h'' \in \mathfrak{H}^q; \end{aligned}$$

therefore $h = \phi^*(g + \sum m_j g_j g_j') + h''$.

Thus $\phi_q{}^*$ is onto. By induction, $\phi_q{}^*$ (and $\psi_q{}^* = (\phi_q{}^*)^{-1}$) is a ring isomorphism for each integer q; since ψ preserves longitude-meridian pairs, so do $\psi_q{}^*$ and $\phi_q{}^*$.

COROLLARY. *$G/G_\omega \approx H/H_\omega$, and $JG/\mathfrak{G}^\omega \approx JH/\mathfrak{H}^\omega$,* where $G_\omega = \bigcap_{q<\infty} G_q$ *and* $\mathfrak{G}^\omega = \bigcap_{q<\infty} \mathfrak{G}^q$.

Proof. The map $\psi_\omega{}^*$ is onto, since $\psi_\omega{}^*\phi_\omega{}^* = 1$. The kernel of the map $JH \to JG/\mathfrak{G}^\omega$ clearly contains $\mathfrak{H}^\omega$, and by the theorem, is contained in $\mathfrak{H}^q$ for every integer q. Therefore $\psi_\omega{}^*$ is an isomorphism. The proof that ψ_ω is an isomorphism is similar.

4. In the manner of Milnor [2], we can find a presentation for the additive group of $JG/\mathfrak{G}^q$, in which the multiplicative structure has a simple form, for a link in the 3-sphere.

Let $1 \to R \to F \to G \to 1$ be a presentation for G. Denoting the fundamental ideal of JF by $\mathfrak{F}$, and the 2-sided ideal generated by elements $\{r - 1 \colon r \in R\}$ by $\mathfrak{R}$, we have $JG/\mathfrak{G}^q \approx JF/\mathfrak{R} \cup \mathfrak{F}^q$.

In particular, we shall take a Wirtinger presentation for G:

$$G = (x_{ij} \colon x_{ij} = u_{ij} x_{i1} u_{ij}^{-1}, \text{ for } i = 1, 2, \cdots, n, j = 1, 2, \cdots, i_m)$$

where u_{ij} is a word in the free group F generated by the x's. In general, u_{i1} is a parallel of the i-th component of L. By multiplication by a suitable power of x_{i1} (since u_{i1} commutes with x_{i1}) we may assume that (u_{i1}, x_{i1}) is an i-th longitude-meridian pair for L.

Write $X_{ij} = x_{ij} - 1$, $U_{ij} = u_{ij} - 1$ in JF. Now $JF/\mathfrak{J}^q$ is the free abelian group on the $2^q - 1$ generators $1, X_{ij}, X_{i_1j_1}X_{i_2j_2}, \cdots, X_{i_1j_1}X_{i_2j_2} \cdots X_{i_{q-1}j_{q-1}}$. The multiplicative structure of $JF/\mathfrak{J}^q$ is given by formal non-commutative multiplication, in which $1 \cdot X_{ij} = X_{ij}$ and products of the X's of length greater than $q - 1$ are identically zero.

The relation ideal $\mathcal{R}$ is additively generated by elements

$$A(u_{ij}x_{i1}u_{ij}^{-1}x_{ij}^{-1} - 1)B, \text{ for all } A, B \in JF.$$

In terms of the generators X_{ij}, we may express these elements as

$$A((1 + U_{ij})(1 + X_{i1}) - (1 + X_{ij})(1 + U_{ij}))B, \text{ or}$$
$$A((U_{ij}X_{i1} - X_{ij}U_{ij}) + (X_{i1} - X_{ij}))B \cdots (R_{ij})$$

Notice that only finitely many multiples A, B need be considered, since products of length q or greater are trivial. Thus $JG/\mathfrak{G}^q$ is finitely presented. The multiplicative structure is induced from that of $JF/\mathfrak{J}^q$

Let $\bar{F}$ be the free group of rank n on generators $x_{11}, x_{21}, \cdots, x_{n1}$.

LEMMA. *Any element $A \in JF$ may be expressed in the form $A = B + C$, where $B \in J\bar{F}$ and $C \in \mathcal{R} \cup \mathfrak{J}^q$, for arbitrary q.*

Proof. We need consider only the case that A is a generator X_{ij}. Since $x_{ij} = u_{ij}x_{i1}u_{ij}^{-1} \bmod \mathcal{R}$,

$$\begin{aligned} X_{ij} &= (1 + U_{ij})(1 + X_{i1})(1 - U_{ij} + U_{ij}^2 - \cdots + (-1)^{q-1}U_{ij}^{q-1}) - 1 \\ &\qquad\qquad \bmod \mathcal{R} \cup \mathfrak{J}^q \\ &= X_{i1} + U_{ij}X_{i1} - X_{i1}U_{ij} + \cdots + (-1)^{q-1}(U_{ij}X_{i1}U_{ij}^{q-3} - X_{i1}U_{ij}^{q-2}) \\ &\qquad\qquad \bmod \mathcal{R} \cup \mathfrak{J}^q \cdots (A_{ij}) \\ &= X_{i1} \bmod \mathcal{R} \cup \mathfrak{J}^2. \end{aligned}$$

Thus the lemma is true for $q = 2$.

Assume the lemma is true for $q < m$. In particular, put $U_{ij}^q = B_q + C_q$, where $B_q \in J\bar{F}$ and $C_q \in \mathfrak{J}^{m-1} \cup \mathcal{R}$, for each $q < m$. Then, in (A_{ij}),

$$\begin{aligned} X_{ij} &= X_{i1} + (B_1 + C_1)X_{i1} - X_{i1}(B_1 + C_1) + \cdots \qquad \bmod \mathcal{R} \cup \mathfrak{J}^m \\ &= X_{i1} + B_1X_{i1} - X_{i1}B_1 + \cdots + (-1)^{m-1}(B_1X_{i1}B_{m-3} - X_{i1}B_{m-2}) \\ &\qquad\qquad \bmod \mathcal{R} \cup \mathfrak{J}^m \end{aligned}$$

since $X_{i1}C_q = C_qX_{i1} = 0 \bmod \mathcal{R} \cup \mathfrak{J}^m$.

The result follows by induction.

This expansion of X_{ij} modulo $\mathcal{R} \cup \mathfrak{J}^q$ in $J\bar{F}$ corresponds with the Chen homomorphism into $\bar{F}$ of [2].

Using this lemma to express the relations in terms of $J\bar{F}$, in particular substituting (A_{ij}) for X_{ij} in (R_{ij}) in the case $j \neq 1$, we find that (R_{ij}) reduces to the trivial relator 0. We are left with the relators (R_{i1}) and the consequences of (A_{ij}). The latter may be used to eliminate the generators X_{ij}, $j \neq 1$; thus, writing X_i for X_{i1}, U_i for U_{i1}, we may simplify the above presentation to

$$JG/\mathcal{G}^q \approx (1, X_i, X_iX_j, X_{i_1}X_{i_2}\cdots X_{i_{q-1}} : A(X_iU_i - U_iX_i)B = 0 \text{ for } A, B \in J\bar{F}).$$

5. We now introduce an invariant of F-isotopy for links of two components in the 3-sphere with non-zero linking number v. In this case, G has a Wirtinger presentation

$$(x_1, \cdots, x_s, y_1, \cdots, y_t : x_i = u_i x_1 u_i^{-1}, y_j = v_j y_1 v_j^{-1}),$$

where the x's are meridians of the first component, and the y's are meridians of the second. Any one of these relations is dependent on the others. Thus we may delete the relation $y_1 = v_1 y_1 v_1^{-1}$. Then, putting $X = x_1 - 1$, $Y = y_1 - 1$, $U = u_1 - 1$, we have

$$JG/\mathcal{G}^q \approx (1, X, Y, XY, YX, \cdots, Y^{q-1} : A(XU - UX)B = 0).$$

Suppose U has Magnus expansion

$$U = a_0Y + h_0X + (a_1XY + bY^2 + cX^2 + dYX) + \cdots$$
$$+ (a_nX^nY + h_nXY^n + \cdots) + \cdots \bmod \mathcal{R} \cup \mathfrak{J}^q$$

where the terms bracketed together are of the same total degree; the reason for singling out the terms X^nY and XY^n will become apparent. Since U is a longitude of the first component, a_0 is the linking number v of L, and $h_0 = 0$; note that $h_1 = a_1$. By Fox [1], the co-efficient of Y^n is the binomial co-efficient ${}^vC_n$. It should be noted that the $\bar{\mu}$-invariants of L are determined by the co-efficients of this expansion.

We define integers a_n^*, h_n^* inductively as follows:

0) $a_0^* = a_0 = v$, and $h_0^* = 0$;

1) a_1^* is that non-negative integer less than v such that $a_1^* \equiv a_1 \bmod v$. Let $c_0 = (a_1^* - a_1)/v$.

h_1^* is that non-negative integer less than v such that $h_1^* \equiv h_1 \bmod v$. Let $k_0 = (h_1^* - h_1)/v$. Of course, $h_1^* = a_1^*$ and $c_0 = k_0$.

n) a_n^* is that non-negative integer less than v such that

$$a_n^* \equiv a_n + \sum_{i=0}^{n-2} c_i a_{n-i-1} \bmod v; \text{ and } c_{n-1} = (a_n^* - a_n - \sum_{i=0}^{n-2} c_i a_{n-i-1})/v.$$

h_n^* is that non-negative integer less than v such that

$$h_n^* \equiv h_n + \sum_{i=0}^{n-2} k_i \cdot {}^vC_{n-i} \bmod v; \text{ and } k_{n-1} = (h_n^* - h_n - \sum_{i=0}^{n-2} k_i \cdot {}^vC_{n-i})/v.$$

These numbers are defined for $n \leqq q - 2$; however, since q may be chosen arbitrarily large, this is no real restriction.

THEOREM. *The numbers* a_n^*, h_n^*, $n \leqq q - 2$, *remain invariant under an F-isotopy of L.*

Proof. By the previous theorem, we must show that a_n^* and h_n^* remain invariant under the following operations:

1) replacing U by $U + R$, where $R \in \mathcal{R} \cup \mathfrak{J}^q$;

2) replacing (u_1, x_1) by a conjugate pair (u_1', x_1'), where $u_1' = au_1a^{-1}$ and $x_1' = ax_1^{-1}$;

3) replacing y_1 by a conjugate $y_1' = by_1b^{-1}$.

1) By Section 4, we need only consider addition of some relation $R = A(XU - UX)B$. The expansion of $XU - UX$ contains no terms of the form X^i or Y^i, but does contain terms $a_{n-1}X^nY$ and ${}^vC_nXY^n$. Thus the only relations which change the co-efficients a_n and h_n are

(i) the relation $XU - UX$, under which a_n becomes $a_n' = a_n + a_{n-1}$, and h_n becomes $h_n' = h_n + {}^vC_n$.

(ii) the relation $X^i(XU - UX)$, under which $a_n' = a_n$ if $i > n$, $a_n' = a_n + a_{n-1}$ if $i \leqq n$, and $h_n' = h_n$.

(iii) the relation $(XU - UX)Y^i$, under which $a_n' = a_n$, $h_n' = h_n + {}^vC_{n-i}$ if $i < n$, and $h_n' = h_n$ if $i \geqq n$.

Then an inductive argument, straightforward but tedious, shows that a_n^* and h_n^* are unaltered by these changes.

2) and 3). Case 2) reduces triviality to case 3) (with $b = a^{-1}$). Let $\langle W : Y^i \rangle$ be the co-efficient of the term W in the Magnus expansion of Y^i in terms of X and $Z = y_1' - 1$. Then by comparing co-efficients we find

a) $\langle Z^j : Y \rangle = 0$ for $j > 1$, $\langle Z : Y \rangle = 1$.

b) $\langle X^j : Y \rangle = 0$ for all j.

c) $\langle Z^j : Y^i \rangle = 0$ for $i < j$.

d) $\langle X^jZ : Y^i\rangle = 0$ for $i > 1$.

e) $\langle XZ : Y\rangle = \langle XZ^n : Y^n\rangle$ for $n \leqq q-2$.

Using these results, we find that a_n becomes

$$a_n' = a_n + \sum_{i=0}^{n-1} \langle X^{i+1}Z : Y\rangle a_{n-i-1},$$

and h_n becomes

$$h_n' = h_n + \sum_{i=0}^{n-1} \langle XZ^n : Y^{n-i}\rangle^{v} C_{n-i}.$$

By e), these changes can also be attained by a combination of the operations (i), (ii), and (iii). Thus again there is no change in a_n^* and h_n^*.

6. Application. Consider the link L_n of the figure. Then a longitude U has expansion

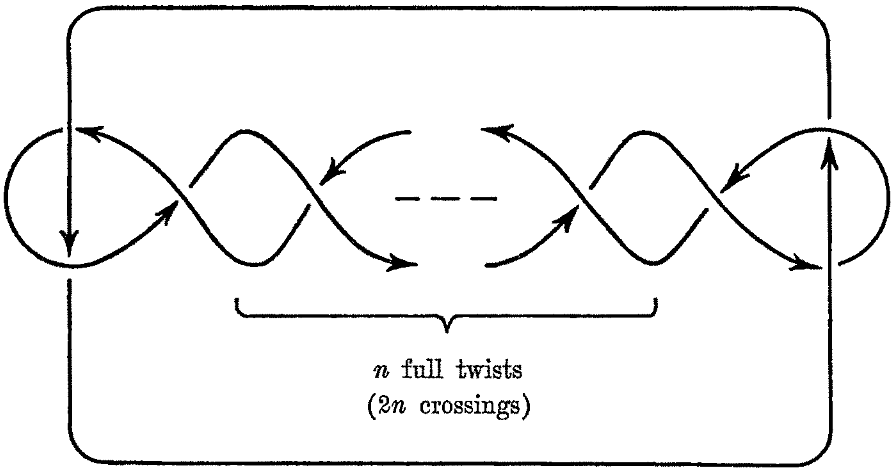
$$U = 2Y + Y^2 + XY - YX - (n+1)Y^2X \\ + (2n+1)YXY - nXY^2 - XYX + YX^2 + \cdots.$$

The Milnor $\bar{\mu}$-invariants are the same for every n. The rings $JG/\mathfrak{G}^q$ appear to be isomorphic for all n; certainly this is true for $q \leqq 5$. However the invariants a_m^*, h_m^* show that L_n is not F-isotopic to $L_{n'}$, for $n \neq n'$.

It is not too difficult to show that for U as above, $a_0 = 2$, $a_1 = 1$, $a_m = 0$ for $m \geqq 2$, and $h_0 = 0$, $h_1 = 1$, $h_2 = -n$, $h_m = 0$ for $m \geqq 3$. n has a unique expansion of the form $n = \sum_{i=0}^{p} r_i 4^i$, where $r_i = -1, 0, 1$ or 2. Define $H(0) = 0$, $H(1) = 3$, $H(2) = 2$, $H(-1) = 1$ and $H(n) = \sum_{i=0}^{p} H(r_i)4^i$.

Then $a_0^* = 2$, $a_1^* = 1$, $a_m^* = 0$ for $m \geqq 2$, and $h_0^* = 0$, $h_1^* = 1$, and h_m^* is the co-efficient of 2^{m-2} in the binary expansion of $H(n)$, for $m \geqq 2$. Clearly these distinguish the F-isotopy classes of the links L_n for every n.

Note that these results hold also for negative values of n; it follows that L_n is not amphicheiral. With any link of two components we may associate 8 oriented links, running through all combinations of orientations for the components and the 3-sphere. Of the 7 other versions of L_n, 4 have linking number -2 and so may be disregarded. Reversing the orientations of the components and leaving the orientation of the 3-sphere fixed clearly gives L_n again (i.e., L_n is invertible). However, the link obtained by reversing the orientation of the 3-sphere and of exactly one component is L_{-n-1}, which we have shown is not F-isotopic to L_n.



L_n

Princeton University.

REFERENCES.

[1] R. H. Fox, "Free differential Calculus I," *Annals of Mathematics*, vol. 57 (1953), pp. 547-560.

[2] J. Milnor, *Isotopy of Links*, *Lefschetz symposium*, Princeton Mathematical Series, vol. 12, pp. 280-306.

[3] N. Smythe, *Isotopy Invariants of Links*, Princeton Ph. D. thesis (1965).

THE HOMOTOPY GROUPS OF WEDGES OF SUSPENSIONS.[1]

By Gerald J. Porter.[2]

In [1] Hilton proved that there exists a set of spheres, $\{S^{r_i}\}_{i=1}^{\infty}$, such that $\pi_m(S^{n_1}\vee\cdots\vee S^{n_k})$ is isomorphic to $\sum_{i=1}^{\infty}\pi_m(S^{r_i})$ for each m, where $S^{n_1}\vee\cdots\vee S^{n_k}$ is the one point union of the spheres, S^{n_i}, $1\leqq i\leqq k$. In fact, Hilton proved that $\Omega(S^{n_1}\vee\cdots\vee S^{n_k})$ is homotopy equivalent to the weak infinite product $\prod_{i=1}^{\infty}\Omega S^{r_i}$.

Let $X_1,\cdots,X_n$ be connected, countable CW-complexes. Milnor [3] generalized Hilton's result to show that $\Omega(\Sigma X_1\vee\cdots\vee\Sigma X_n)$ is homotopy equivalent to a weak infinite product, $\prod_{i=1}^{\infty}\Omega\Sigma Y_i$, where each Y_i equals $\wedge(X_1^{(i_1)},\cdots,X_n^{(i_n)})$ for some n-tuple, $(i_1,\cdots,i_n)$, and $\wedge(X_1^{(i_1)},\cdots,X_n^{(i_n)})$ is the smash product of i_j copies of X_j, $1\leqq j\leqq n$. Moreover, an explicit algorithm is given for computing the set of n-tuples over which the product is taken.

At the other extreme, it is trivially shown that $\Omega(\Sigma X_1\times\cdots\times\Sigma X_n) = \prod_{i=1}^{n}\Omega\Sigma X_i$. The principal result of this paper extends the Hilton-Milnor theorem to those wedges lying between the one point union and the cartesian product.

In particular, for $0\leqq i<n$, set $T_i(A_1,\cdots,A_n)$ equal to the subset of $A_1\times\cdots\times A_n$ with at least i coordinates at base points. T_0 is then the cartesian product and T_{n-1} is the one point union.

For each i, $0\leqq i<n$, it is proven that $\Omega T_i(\Sigma X_1,\cdots,\Sigma X_n)$ is homotopy equivalent to the weak infinite product, $\prod_{j=1}^{\infty}\Omega\Sigma Y_j$, where each

$$Y_j=\Sigma^k\wedge(X_1^{(j_1)},\cdots,X_n^{(j_n)})$$

for some $(n+1)$-tuple, $(k,j_1,\cdots,j_n)$. In addition, the set of $(n+1)$-

---

Received January 11, 1965.
Revised July 8, 1965.

[1] The results of this paper were announced in [6].

[2] This research was supported in part by National Science Foundation Grant GP-1740.

tuples over which the product is taken is easily computed and depends only upon i and n.

Since $T_{n-i}(\Sigma X_1,\cdots,\Sigma X_n)$ is the universal example for a certain type of i-ary homotopy operation on n variables, this result may be interpreted as a classification theorem for such operations.

In an earlier paper [5] this result was proven for $T_1(S^{n_1},\cdots,S^{n_k})$. The methods used in that paper were highly algebraic and relied upon a knowledge of the ring $H_*(\Omega T_1)$. The methods used here are much more geometric.

By replacing the inclusion, $j\colon T_i \to T_0$, with a homotopy equivalent fibre map, p, it is shown that $\Omega T_i(X_1,\cdots,X_n)$ is homotopy equivalent to $\Omega X_1 \times \cdots \times \Omega X_n \times \Omega F_i$, where F_i is the fibre of p. A homotopy analysis then proves that F_i is homotopy equivalent to a one point union of spaces of the form, $\Sigma^r \wedge (\Omega X_{j_1},\cdots,\Omega X_{j_k})$ where $1 \leqq j_1 < \cdots < j_k \leqq n$, and $r \geqq 1$.

Hence by the Hilton-Milnor theorem, ΩF_i is homotopy equivalent to the cartesian product of spaces of the form $\Omega\Sigma^k \wedge ((\Omega X_1)^{(j_1)},\cdots,(\Omega X_n)^{(j_n)})$. If each $X_i = \Sigma Y_i$, Milnor's methods may be employed to prove that each of the above spaces is homotopy equivalent to the weak infinite product of spaces of the form $\Omega\Sigma^t \wedge (Y_1^{(k_1)},\cdots,Y_n^{(k_n)})$. This combined with the above result completes the proof.

Section 1. *Notation.* It will be assumed throughout that all spaces have the homotopy type of a *connected,* countable CW-complex with base point and that all maps are continuous and base point preserving. The category of these spaces and maps is denoted by $\mathcal{C}$. $\mathcal{C}^n$ is the category of n-tuples of objects and maps of $\mathcal{C}$.

The functors $T_i\colon \mathcal{C}^n \to \mathcal{C}$, $i = 0, 1, \cdots, n-1$, are defined by setting $T_i(X_1,\cdots,X_n)$ equal to the subset of the cartesian product, $X_1 \times \cdots \times X_n$, consisting of those points with at least i coordinates at base points. The one point union, $T_{n-1}(X_1,\cdots,X_n)$, will also be denoted by $\bigvee_{j=1}^{n} X_j$.

The smash product functor, $\wedge\colon \mathcal{C}^n \to \mathcal{C}$, is the quotient functor, T_0/T_1. Using this, define $X^{(n)}$ inductively as follows: $X^{(0)} = S^0$, $X^{(n)} = \wedge(X^{(n-1)}, X)$. The n-fold suspension, $\Sigma^n\colon \mathcal{C} \to \mathcal{C}$, is defined to be $\wedge(S^n,\ )$.

The path space of X, $P(X)$, is the set of maps, $f\colon I \to X$, such that $f(0) = *$. The loop space of X, ΩX, is the subset of $P(X)$ consisting of those maps, f, with $f(1) = *$.

Σ, Ω, and P induce functors $\mathcal{C}^n \to \mathcal{C}^n$ denoted by the same symbols. In particular $\Sigma(X_1,\cdots,X_n)$ equals $(\Sigma X_1,\cdots,\Sigma X_n)$. Ω and P are defined similarly.

The Basic Splitting. Let $j\colon T_i \to T_0$ be the inclusion map. It is well known that j induces a homotopy equivalent fibre map p, giving the following diagram:

$$\begin{tikzcd}
 & T_i \arrow[r, "j"] \arrow[d, "h"] & T_0 \arrow[d, equal] \\
F_i \arrow[r, "u"] & E \arrow[r, "p"] & T_0
\end{tikzcd}$$

where i) $E = \{(x,\eta) \epsilon T_i \times T_0{}^I \mid j(x) = \eta(1)\}$
ii) $p(x,\eta) = \eta(0)$
iii) $F_i = p^{-1}(*)$
iv) $u =$ inclusion
v) $h(x) = (x,\eta_x)$ where η_x is the constant path at x.
vi) h is a homotopy equivalence
vii) $ph = j$.

LEMMA 1. *There exists $\theta\colon \Omega T_0(X_1,\cdots,X_n) \to \Omega E$ such that $(\Omega p)\theta$ is homotopic to the identity. Moreover, if $n - i \geqq 2$, θ can be chosen to be an H-homomorphism.*

Proof. (See also 2.14 of [2]) Since $ph = j$, it suffices to find $\mu\colon \Omega T_0 \to \Omega T_i$, such that $(\Omega j)\mu \sim 1$. Let $m_k\colon \Omega X_k \to \Omega T_i$, $k = 1,\cdots,n$, be induced by the injection $X_k \to T_i$. Set μ equal to the product, $m_1 \cdot \cdots \cdot m_n$, in ΩT_i. When $n - i \geqq 2$, the m_k pairwise homotopy commute in ΩT_i and thus μ is an H-homomorphism. Clearly $(\Omega j)\mu \sim 1$.

An adaptation of the standard group theory argument can be used to prove the following lemma.

LEMMA 2. *Let $* \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to *$ be a short exact sequence of H-spaces and H-homomorphisms. If $\gamma\colon C \to B$ is such that $\beta\gamma \sim 1$, then $B \sim A \times C$ (not necessarily as H-spaces).*

Since Ω preserves fibrations, the following sequence is exact.

$$* \to \Omega F_i \xrightarrow{\Omega u} \Omega E \xrightarrow{\Omega p} \Omega T_0 \to *.$$

Hence Lemma 1 and Lemma 2 imply that if $X_1,\cdots,X_n$ are simply connected then:

COROLLARY. $\Omega T_i(X_1,\cdots,X_n) \sim \Omega E \sim \Omega X_1 \times \cdots \times \Omega X_n \times \Omega F_i$.

Thus the study of the homotopy type of $\Omega T_i(X_1,\cdots,X_n)$ is reduced to a study of the homotopy type of ΩF_i.

Section 2. The following theorem will be proven in Section 3.

THEOREM 1. *F_i is homotopy equivalent to*

$$\bigvee_S \overset{r}{\bigvee} \Sigma^{n-i} \wedge \Omega(X_{k_1},\cdots,X_{k_j})$$

with $S=\{(k_1,\cdots,k_j) \mid 1 \leqq k_1 < \cdots < k_j \leqq n$ and $n-i+1 \leqq j \leqq n\}$ and with r equal to the binomial coefficient $\binom{j-1}{n-i}$, where $\overset{r}{\bigvee} X$ is the one point union of r copies of X.

When $i=1$, F_i has a particularly simple form; namely,

$$F_1 \sim \Sigma^{n-1} \wedge \Omega(X_1,\cdots,X_n),$$

and hence

$$\Omega T_1(X_1,\cdots,X_n) \sim T_0(\Omega X_1,\cdots,\Omega X_n,\Omega\Sigma^{n-1} \wedge \Omega(X_1,\cdots,X_n)).$$

As an example let $X_1=X_2=\cdots=X_n=CP^\infty$. Since $\Omega CP^\infty = S^1$,

$$\Omega T_1(CP^\infty,\cdots,CP^\infty) \sim T_0(S^1,\cdots,S^1,\Omega S^{2n-1}).$$

$\bigvee$-Basic Products. Milnor [3] proved[3] that $\Omega(\bigvee_{i=1}^{k} \Sigma A_i)$ is homotopy equivalent to a weak infinite product, $\prod_{i=1}^{\infty} \Omega\Sigma A_i$. The set of such A_i is called a set of $\bigvee$-basic products and is defined inductively as follows. The products of weight one are $A_1,\cdots,A_k$. A function $e(i)$ is defined for $i \leqq k$ by $e(i)=0$. Assume that the products of weight less than m have been defined and ordered and that $e(i)$ is defined for such A_i. The products of weight m are the set of all $\wedge(A_i,A_j)$ such that weight $A_i +$ weight $A_j = m$ and $e(i) \leqq j < i$. These are ordered arbitrarily among themselves and are greater than all products of lesser weight. Set $e(h)=j$ if $A_h = \wedge(A_i,A_j)$. This completes the description of the set of $\bigvee$-basic products.

If each A_i, for $i \leqq k$, is equal to $\Sigma^t \wedge ((\Omega X_1)^{(j_1)},\cdots,(\Omega X_n)^{(j_n)})$ for some $(n+1)$-tuple, $(t,j_1,\cdots,j_n)$, it is easily seen that when $i > k$, each basic product, A_i, is also of this form. Hence Milnor's theorem implies:

THEOREM 2. *ΩF_i is homotopy equivalent to a weak infinite product of spaces of the form $\Omega\Sigma^t \wedge ((\Omega X_1)^{(j_1)},\cdots,(\Omega X_n)^{(j_n)})$. In addition the set of $(n+1)$-tuples, $(t,j_1,\cdots,j_n)$, over which the product is taken, is computable.*

---

[3] An alternate proof can be given using the methods of this paper.

In general it is impossible to obtain a simpler form for ΩF_i than that given by Theorem 2. However, if each $X_j = \Sigma Y_j$ for some $Y_j \in \mathscr{C}$, the decomposition of ΩF_i may be continued as is seen by the following theorem.

THEOREM 3. *$\Omega\Sigma^t \wedge \Omega\Sigma(Y_1, \cdots, Y_n)$ is homotopy equivalent to a weak infinite product, $\prod_{j=n+1}^{\infty} \Omega\Sigma Y_j$, if $t \geqq 1$, where each $Y_j = \Sigma^k \wedge (Y_1{}^{(j_1)}, \cdots, Y_n{}^{(j_n)})$ for some $(n+1)$-tuple, $(k, j_1, \cdots, j_n)$.*

$\wedge$-Basic Products. The set of Y_j in the preceding theorem are called a set of $\wedge$-basic products and are defined inductively as follows. The basic products of weight one are $Y_1, \cdots, Y_n$, and $\Sigma^{t-1} \wedge (Y_1, \cdots, Y_n) = Y_{n+1}$. Those of weight two are $Y_{n+1+j} = \wedge (Y_{n+1}, Y_j)$ for $j = 1, \cdots, n$. Define e by setting $e(h) = 0$ if $1 \leqq h \leqq n+1$ and $e(h) = h - (n+1)$ if $n+1 < h \leqq 2n+1$. Assume the products of weight less than m have been defined and are ordered for $m > 2$ and that $e(i)$ is defined for all such i. The basic products of weight m are all elements $\wedge (Y_i, Y_j)$ such that weight $Y_i +$ weight $Y_j = m$ and $e(i) \leqq j < i$. These are ordered arbitrarily among themselves and are greater than all products of lesser weight. Set $e(h) = j$ if $Y_h = \wedge (Y_i, Y_j)$. This completes the definition of the set of $\wedge$-basic products.

The proof of Theorem 3 will follow from several lemmas. The following two lemmas were proven in [3]. Indeed, the proof of Theorem 3 is modeled after [3].

LEMMA 3. $\Omega(\Sigma X \vee \Sigma Y) \sim \Omega\Sigma X \times \Omega\Sigma(\bigvee_{i=0}^{\infty} \wedge (Y, X^{(i)}))$.

LEMMA 4. $\Sigma\Omega\Sigma X \sim \bigvee_{i=1}^{\infty} \Sigma X^{(i)}$.

For each m, set $R_m = \bigvee_{\substack{h \geqq m \\ e(h) < m}} \Sigma Y_h$, where $\{Y_h\}$ is a set of $\wedge$-basic products.

LEMMA 5. *$R_{n+1} \sim \Sigma^t \wedge \Omega\Sigma(Y_1, \cdots, Y_n)$ if $t \geqq 1$.*

Proof. Since $t \geqq 1$, the distributivity of Σ under $\wedge$ and Lemma 4 imply that

$$\begin{aligned}\Sigma^t \wedge \Omega\Sigma(Y_1, \cdots, Y_n) &\sim \Sigma^t \wedge (\bigvee_{k_1=1}^{\infty} Y_1{}^{(k_1)}, \cdots, \bigvee_{k_n=1}^{\infty} Y_n{}^{(k_n)}) \\ &= \bigvee_{(j_1, \cdots, j_n)} \Sigma \wedge (Y_{n+1}, Y_1{}^{(j_1)}, \cdots, Y_n{}^{(j_n)})\end{aligned}$$

where $(j_1, \cdots, j_n)$ runs over all n-tuples of non-negative integers. Writing $\wedge (Y_{n+1}, Y_1{}^{(j_1)}, \cdots, Y_n{}^{(j_n)})$ as

$$\wedge (\cdots \wedge (\wedge (Y_{n+1}, Y_1), \cdots), Y_r)$$

(where Y_i appears j_i times for $i \leqq r < n$ and $j_i = 0$ for $i > r$) it is easily seen that each such term is a $\wedge$-basic product with $h \geqq n+1$ and $e(h) < n+1$. Thus it suffices to show that if Y_h is a $\wedge$-basic product with $h \geqq n+1$ and $e(h) < n+1$ it has the above form.

This is done by induction on $e(h)$. If $e(h) = 0$, h must equal $n+1$ and $Y_h = \wedge(Y_{n+1}, Y_1^{(0)}, \cdots, Y_n^{(0)})$. Assume inductively that Y_h has the above form whenever $h \geqq n+1$ and $e(h) < k \leqq n$. Let $e(h) = k$. Then $Y_h = \wedge(Y_s, Y_k)$ and $e(s) \leqq k$. If $e(s) = k$, we repeat the above process. Since each basic product is the smash product of a finite number of Y_i,

$$Y_h = \wedge(\cdots(\wedge(Y_r, Y_k), \cdots), Y_k)$$

where $e(r) < k$. Thus using our inductive hypothesis on Y_r the lemma is proven.

LEMMA 6. *$\Omega\Sigma Y_m \times \Omega R_{m+1} \sim \Omega R_m$ if $m > n$.*

Proof. $R_m = \Omega(\Sigma Y_m \vee \Sigma Z)$ where $Z = \bigvee_{h>m} Y_h$. By Lemma 3:

$$\begin{aligned}\Omega R_m &\sim \Omega\Sigma Y_m \times \Omega\Sigma\Big(\bigvee_{i=0}^{\infty} (Z, Y_m^{(i)})\Big)\\ &= \Omega\Sigma Y_m \times \Omega\Sigma\Big(\bigvee_{i=0}^{\infty} \bigvee_{\substack{h>m\\ e(h)<m}} \wedge(Y_h, Y_m^{(i)})\Big)\\ &= \Omega\Sigma Y_m \times \Omega\Big(\bigvee_{\substack{h\geqq m+1\\ e(h)<m+1}} \Sigma Y_h\Big)\\ &= \Omega\Sigma Y_m \times \Omega R_{m+1}.\end{aligned}$$

COROLLARY. $\Omega R_{n+1} \sim \prod_{i=n+1}^{\infty} \Omega\Sigma Y_i$.

Lemma 5 and the above corollary complete the proof of Theorem 3.

Combining Theorems 1, 2, and 3 with the Hilton-Milnor theorem we have the following theorem:

THEOREM 4. *$\Omega T_l\Sigma(X_1, \cdots, X_n)$ is homotopy equivalent to a weak infinite product $\prod_{j=1}^{\infty} \Omega\Sigma A_j$ where each A_j equals $\Sigma^k \wedge (X_1^{(j_1)}, \cdots, X_n^{(j_n)})$ for some $(n+1)$-tuple, $(k, j_1, \cdots, j_n)$. In addition, the set of $(n+1)$-tuples over which the product is taken, is computable.*

As an example let $T_5^{10}(S^n) = T_5(A_1, \cdots, A_{10})$ where $A_1 = \cdots = A_{10} = S^n$. If $r < 11n - 6$

$$\pi_r(T_5{}^{10}(S^n)) = \sum_{j=1}^{10} \pi_r(S^n) \oplus \sum_{k=6}^{10} \sum_{j=1}^{i_k} \pi_r(S^{k(n-1)+5})$$

where $\sum_{j=1}^{n} G = G \oplus \cdots \oplus G$ and $i_6 = 210$, $i_7 = 1980$, $i_8 = 10{,}395$, $i_9 = 40{,}040$ and $i_{10} = 126{,}126$.

Section 3. The homotopy theory necessary to prove Theorem 1 is developed in this section. A more general result is proven and Theorem 1 is deduced as a corollary. The author wishes to thank Professor George Whitehead for several helpful discussions.

The fibre F_i, of Theorem 1, is easily seen to be the set of maps, $f\colon I \to T_0(X_1, \cdots, X_n)$, such that

$$f(0) = (*, \cdots, *) \text{ and } f(1) \in T_i(X_1, \cdots, X_n).$$

This is the subspace of $T_0P(X_1, \cdots, X_n)$ consisting of those n-tuples in which at least i of the paths are loops. ΩX_i is a subcomplex of the contractible space PX_i and hence Theorem 1 follows from the following theorem.

THEOREM 5. *For each i, $1 \leqq i \leqq n$, let K_i be a contractible space containing X_i as a subspace (i.e. subcomplex). Set $J_k(X_1, \cdots, X_n)$ equal to the subset of $T_0(K_1, \cdots, K_n)$ consisting of those n-tuples, $(y_1, \cdots, y_n)$, such that $y_i \in X_i$ for at least k values of i. If $k < n$, $J_k(X_1, \cdots, X_n)$ is homotopy equivalent to*

$$\bigvee_S \overset{r}{\bigvee} \Sigma^{n-k} \wedge (X_{i_1}, \cdots, X_{i_j})$$

with $S = \{(i_1, \cdots, i_j) \mid 1 < \cdots < i_j \leqq n$ and $n - k + 1 \leqq j \leqq n\}$ and with r equal to the binomial coefficient $\begin{pmatrix} j-1 \\ n-k \end{pmatrix}$.

J_k is a generalization of the join of the spaces, $X_1, \cdots, X_n$. In particular, $J_1(X_1, \cdots, X_n)$ is the usual join. In this case, Theorem 5 follows from Theorems 2 and 4 of the appendix of [4].

Theorem 5 will be proven by induction on n. Several lemmas are needed to prepare for the induction.

LEMMA 7. *Let A and B be contractible as subcomplexes of X and Y, respectively. Then $X \times B \underset{A \times B}{\bigcup} A \times Y$ is homotopy equivalent to*

$$((X \times B)/B) \vee (\Sigma \wedge (A, B)) \vee ((A \times Y)/A).$$

Proof. Since A is contractible in X, there is a map, $h\colon CA \to X$, where CA, the cone of A, is $\wedge(I, A)$. (The basepoint of I is taken to be 0.) The mapping cylinder of h is homotopy equivalent to X and contains CA as a subcomplex. Thus, if necessary by replacing A by the mapping cylinder of h, it may be assumed that $A \subset CA \subset X$ and similarly $B \subset CB \subset Y$. Therefore,

$$X \times B \bigcup_{A\times B} A \times Y = X \times B \bigcup_{CA\times B} (CA \times B \cup A \times CB) \bigcup_{A\times CB} A \times Y.$$

The proof of the lemma is completed by factoring out the contractible subspace $CA \vee CB$ and noting that $\wedge(CA, Y)$ and $\wedge(X, CB)$ are contractible.

COROLLARY. *If X and Y are contractible, $X \times B \bigcup_{A\times B} A \times Y$ is homotopy equivalent to $\Sigma \wedge (A, B)$.*

LEMMA 8. *$J_k(X_1, \cdots, X_n)$ is contractible in $J_{k-1}(X_1, \cdots, X_n)$ for $0 < k \leqq n$.*

Proof. If $n = 2$, first contract X_1 in K_1 and then contract X_2 in K_2. Assume inductively that the lemma is proven for $n - 1$. Abbreviate $J_k(X_1, \cdots, X_n)$ by $J_k{}^n$. Then

$$J_k{}^n = J_{k-1}{}^{n-1} \times X_n \cup J_k{}^{n-1} \times K_n.$$

Using the inductive hypothesis and Lemma 7

$$J_k{}^n \sim \Sigma \wedge (J_k{}^{n-1}, X_n) \vee (J_{k-1}{}^{n-1} \times X_n)/X_n$$

and similarly

$$J_{k-1}{}^n \sim \Sigma \wedge (J_{k-1}{}^{n-1}, X_n) \vee (J_{k-2}{}^{n-1} \times X_n)/X_n.$$

Since $J_k{}^{n-1}$ is contractible in $J_{k-1}{}^{n-1}$ and $J_{k-1}{}^{n-1}$ is contractible in $J_{k-2}{}^{n-1}$, it follows that $\Sigma \wedge (J_k{}^{n-1}, X_n)$ is contractible in $\Sigma \wedge (J_{k-1}{}^{n-1}, X_n)$ and $(J_{k-1}{}^{n-1} \times X_n)/X_n$ is contractible in $(J_{k-2}{}^{n-1} \times X_n)/X_n$, which proves the lemma.

COROLLARY. *$J_k{}^n$ is homotopy equivalent to*

$$\Sigma \wedge (J_k{}^{n-1}, X_n) \vee (J_{k-1}{}^{n-1} \times X_n)/X_n \text{ for all } n \text{ with } 0 < k \leqq n.$$

Since we may assume inductively that $J_{k-1}{}^{n-1}$ is a suspension, it remains to describe spaces of the form $(X \times \Sigma Y)/X$. This is done in the following lemma.

LEMMA 9. *$(X \times \Sigma Y)/X$ is homotopy equivalent to $\Sigma \wedge (X, Y) \vee \Sigma Y$.*

Proof. Lemma 7.2 of [7] and the J. H. C. Whitehead Theorem imply that $\Sigma(X \times Y) \sim \Sigma X \vee \Sigma Y \vee \Sigma \wedge (X, Y)$. Lemma 9 follows by noting that

$$(X \times \Sigma Y)/X = \Sigma((X \times Y)/X) = \Sigma(X \times Y)/\Sigma X.$$

COROLLARY. $J_k{}^n \sim \Sigma \wedge (J_k{}^{n-1}, X_n) \vee J_{k-1}{}^{n-1} \vee \wedge (J_{k-1}{}^{n-1}, X_n)$.

Finally, one additional lemma is needed.

LEMMA 10. $\Sigma \wedge (J_{n-1}{}^{n-1}, X_n) \sim \vee_S \Sigma \wedge (X_{i_1}, \cdots, X_{i_j}, X_n)$, *where*

$$S = \{(i_1, \cdots, i_j) \mid 1 \leqq i_1 < \cdots < i_j \leqq n-1 \text{ and } 1 \leqq j \leqq n-1\}.$$

Proof. As above, $\Sigma(X \times Y) \sim \Sigma X \vee \Sigma Y \vee \Sigma \wedge (X, Y)$. Iterating this result,

$$\Sigma T_0(X_1, \cdots, X_{n-1}) \sim \vee_S \Sigma \wedge (X_{i_1}, \cdots, X_{i_j}, S^0)$$

from which Lemma 10 follows.

Theorem 5 is proven by induction. The inductive step follows from the Corollary to Lemma 9, Lemma 10, and the fact that $J_0{}^n$ is contractible. The actual calculation is left to the reader.

We conclude by conjecturing that Theorem 4 also holds for the 'non-homogeneous' wedges. A difficulty in these cases seem to lie in the proof of Lemma 8.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.
NOW at: UNIVERSITY OF PENNSYLVANIA.

REFERENCES.

[1] P. J. Hilton, "On the homotopy groups of the union of spheres," *Journal of the London Mathematical Society*, vol. 30 (1955), pp. 154-172.

[2] ———, "Nilpotency and *H*-spaces," *Topology*, vol. 3 (Supplement) (1965), pp. 161-176.

[3] J. Milnor, *The construction FK*, Princeton University (1956) mimeographed.

[4] G. J. Porter, "Higher order Whitehead products," *Topology*, vol. 3 (1965), pp. 123-135.

[5] ———, "On the homotopy groups of the wedge of spheres," *American Journal of Mathematics*, vol. 87 (1965), pp. 297-314.

[6] ———, "A generalization of the Hilton-Milnor Theorem," *Bulletin of the American Mathematical Society*, vol. 71 (1965), pp. 357-359.

[7] G. W. Whitehead, "Homotopy groups of joins and unions," *Transactions of the American Mathematical Society*, vol. 83 (1956), pp. 55-69.

SOME THEOREMS CONCERNING PARTITIONS INTO ODD SUMMANDS.

By Peter Hagis, Jr.

1. Introduction. Let n and t be positive integers, and let $q(n,t)$ represent the number of partitions of n into odd summands, no summand appearing more than t times. Thus, $q(n,1)$ is the number of partitions of n into odd and unequal parts. In [2] I obtained a convergent series and asymptotic formulae for $q(n,1)$ by using the refined circle dissection method of Rademacher [5]. The purpose of the present paper is to generalize these results and obtain a convergent series representation and asymptotic formulae for $q(n,t)$ subject only to the obvious restriction that $n \geqq t$.

The application of the circle dissection method requires first that transformation equations be derived which exhibit the behavior of the generating function near its singularities at the rational points (those with argument $2\pi h/k$) of the unit circle. In [1] such equations were established for the generating function of $q(n)$, the number of partitions of n into odd summands. Using a simple identity relating the generating functions of $q(n)$ and $q(n,1)$ these transformation formulae were exploited in [2] to obtain similar equations for the generating function of $q(n,1)$. In Section 2 of the present paper we shall again make use of the transformation equations of [1], and by employing a simple identity connecting the generating functions of $q(n)$ and $q(n,t)$ shall derive the transformation equations pertinent to the study of the latter.

The Rademacher method also requires that the magnitude of certain complicated sums of roots of unity be estimated rather sharply. In [1] I established the estimates of the sums associated with $q(n)$ by using essentially the same methods employed by Lehner [4]. The sums associated with $q(n,1)$ were estimated in [2] with the argument resting on identities established in [1]. These identities will be used again in Section 3 of the present paper to establish the necessary estimates for the exponential sums arising in the study of $q(n,t)$. We remark that these estimates could be sharpened by using the results on Kloosterman sums to be found in [7], but this is not necessary for our purpose.

Received April 22, 1965.

In Section 4 the desired series for $q(n,t)$ will be derived by making use of the results established in Sections 2 and 3. Several special cases will be considered in Section 5. Finally, in Section 6, asymptotic formulae will be obtained.

2. The transformation equations. With $q(0)=1$ and $|x|<1$ we see that the generating function of $q(n)$ is

$$F(x)=\prod_{m=0}^{\infty}(1-x^{2m+1})^{-1}=\sum_{n=0}^{\infty}q(n)x^{n}.$$

The generating function of $q(n,t)$ is easily seen to be

$$F(x,t)=\prod_{m=0}^{\infty}(1+x^{2m+1}+x^{4m+2}+\cdots+x^{2tm+t})$$

$$=F(x)/F(x^{t+1})=\sum_{n=0}^{\infty}q(n,t)x^{n}.$$

Here $|x|<1$, and we have defined $q(0,t)=1$. Finally, for $|x|<1$ we define

$$H(x)=1/F(x)=\prod_{m=1}^{\infty}(1+x^{m})^{-1}=\sum_{n=0}^{\infty}c(n)x^{n}.$$

Now let h and k be relatively prime integers with $k\geqq 1$, and let $\mathcal{R}(z)>0$. Then Theorem 1 in [1] states that

$$F(\exp\{2\pi ih/k-2\pi z/k\})=\omega(h,k)\exp\{\pi(z-1/z)/12k\}\cdot F(\exp\{2\pi ih'/k-2\pi/zk\}) \tag{2.1}$$

if $2\mid k$, and

$$F(\exp\{2\pi ih/k-2\pi z/k\})=2^{-\frac{1}{2}}\chi(h,k)\exp\{\pi(2z+1/z)/24k\}\cdot H(\exp\{2\pi iH'/k-\pi/zk\}) \tag{2.2}$$

if $2\nmid k$.

Here $hh'\equiv-1\pmod{k}$ and $2hH'\equiv-1\pmod{k}$.

$$\omega(h,k)=\exp\{\pi i\textstyle\sum((\mu/k))((h\mu/k))\},\quad \mu=1,3,\cdots,k-1; \tag{2.3}$$

$$\chi(h,k)=\exp\{\pi i\textstyle\sum((\mu/2k))((h\mu/k))\},\quad \mu=1,3,\cdots,2k-1; \tag{2.4}$$

where $((x))=0$ if x is an integer and $((x))=x-[x]-\frac{1}{2}$ otherwise.

With the aid of this theorem we can now develop a transformation theory for $F(x,t)$. We shall consistently write $D=(k,t+1)$, $k=Dk^{*}$, $t+1=DT$ where, of course, $(k^{*},T)=1$. If we now take

$$x=\exp\{2\pi ih/k-2\pi z/k\} \tag{2.5}$$

then $x^{t+1} = \exp\{2\pi Th/k^* - 2\pi Tz/k^*\}$.

Since $F(x,t) = F(x)/F(x^{t+1})$ we have, if $2 \mid k^*$, by applying (2.1)

$$F(x,t) = w(h,k,t)\exp\{\pi(-tz + (D-T)/Tz)/12k\}$$
$$\cdot F(\exp\{2\pi ih'/k - 2\pi/zk\})H(\exp\{2\pi ih^*/k^* - 2\pi/Tzk^*\}).$$

Here $Thh^* \equiv -1 \pmod{k^*}$, and

$$w(h,k,t) = \omega(h,k)/\omega(Th,k^*). \tag{2.6}$$

If we now define T' as a fixed solution of the congruence $TT' \equiv 1 \pmod{k^*}$ and let

$$y = \exp\{2\pi iT'h'/k - 2\pi/Tzk\}, \tag{2.7}$$

then we easily verify that $\exp\{2\pi ih^*/k^* - 2\pi/Tzk^*\} = y^D$. Also,

$$\exp\{2\pi ih'/k - 2\pi/zk\} = y^T \exp\{2\pi ih'(1-TT')/k\}.$$

Since $1 - TT' = mk^*$ where m is an integer we have, if $m \equiv d \pmod D$ and $h' \equiv b \pmod D$, $\exp\{2\pi ih'/k - 2\pi/zk\} = C(b,d,D)y^T$ where $C(b,d,D) = \exp\{2\pi ibd/D\}$. Therefore, if we define

$$\begin{aligned} G(y,t,k,h) = F(Cy^T)H(y^D) &= \sum_{n=0}^{\infty} q(n)C^n y^{Tn} \sum_{n=0}^{\infty} c(n)y^{Dn} \\ &= \sum_{n=0}^{\infty} c(n,b,d,t,k)y^n \end{aligned} \tag{2.8}$$

we have, for the case $2 \mid k^*$,

$$F(x,t) = w(h,k,t)\exp\{\pi(-tz + (D-T)/Tz)/12k\}G(y,t,k,h).$$

If $2 \mid k$ and $2 \nmid k^*$ we have from (2.1) and (2.2)

$$F(x,t) = 2^{\frac{1}{2}}W(h,k,t)\exp\{\pi(-tz - (2T+D)/2Tz)/12k\}$$
$$\cdot F(\exp\{2\pi ih'/k - 2\pi/zk\})F(\exp\{2\pi iH^*/k^* - \pi/Tzk^*\}).$$

Here $2ThH^* \equiv -1 \pmod{k^*}$, and

$$W(h,k,t) = \omega(h,k)/\chi(Th,k^*). \tag{2.9}$$

If we define T^* as a fixed solution of the congruence $2TT^* \equiv 1 \pmod{k^*}$ and let

$$Y = \exp\{2\pi iT^*h'/k - \pi/Tzk\}, \tag{2.10}$$

then we easily verify that $\exp\{2\pi iH^*/k^* - \pi/Tzk^*\} = Y^D$, and

$$\exp\{2\pi ih'/k - 2\pi/zk\} = Y^{2T}\exp\{2\pi ih'(1-2TT^*)/k\}.$$

Since $1 - 2TT^* = mk^*$ we have, if $m \equiv d (\operatorname{mod} D)$ and $h' \equiv b (\operatorname{mod} D)$, $\exp\{2\pi i h'/k - 2\pi/zk\} = C(b, d, D) Y^{2T}$ where $C(b, d, D) = \exp\{2\pi i bd/D\}$. Therefore, if we define

$$(2.11) \quad H(Y, t, k, h) = F(CY^{2T}) F(Y^D) = \sum_{n=0}^{\infty} q(n) C^n Y^{2Tn} \sum_{n=0}^{\infty} q(n) y^{Dn}$$

$$= \sum_{n=0}^{\infty} d(n, b, d, t, k) Y^n$$

we have, if $2 \mid k$ and $2 \nmid k^*$,

$$F(x, t) = 2^{\frac{1}{2}} W(h, k, t) \exp\{\pi(-tz - (2T + D)/2Tz)/12k\} H(Y, t, k, h).$$

If $2 \nmid k$ we have from (2.2)

$$F(x, t) = z(h, k, t) \exp\{\pi(-tz + (T - D)/2Tz)/12k\}$$
$$\cdot H(\exp\{2\pi i H'/k - \pi/zk\}) F(\exp\{2\pi i H^*/k^* - \pi/Tzk^*\}),$$

where $2hH' \equiv -1 (\operatorname{mod} k)$, $2ThH^* \equiv -1 (\operatorname{mod} k^*)$ and

$$(2.12) \qquad z(h, k, t) = \chi(h, k)/\chi(Th, k^*).$$

With T' a fixed solution of $TT' \equiv 1 (\operatorname{mod} k^*)$, and

$$(2.13) \qquad X = \exp\{2\pi i T'H'/k - \pi/Tzk\}$$

we have $\exp\{2\pi i H^*/k^* - \pi/Tzk^*\} = X^D$. Also,

$$\exp\{2\pi i H'/k - \pi/zk\} = X^T \exp\{2\pi i H'(1 - TT')/k\}.$$

Since $1 - TT' = mk^*$ we have, if $m \equiv d (\operatorname{mod} D)$ and $H' \equiv b (\operatorname{mod} D)$, $\exp\{2\pi i H'/k - \pi/zk\} = C(b, d, D) X^T$. Therefore, if we define

$$(2.14) \qquad J(X, t, k, h) = H(CX^T) F(X^D) = \sum_{n=0}^{\infty} c(n) C^n X^{Tn} \sum_{n=0}^{\infty} q(n) X^{Dn}$$

$$= \sum_{n=0}^{\infty} e(n, b, d, t, k) X^n$$

we have, if $2 \nmid k$,

$$F(x, t) = z(h, k, t) \exp\{\pi(-tz + (T - D)/2Tz)/12k\} J(X, t, k, h).$$

With x, y, Y, X given by (2.5), (2.7), (2.10), (2.13) respectively we combine these results to obtain

THEOREM 1. *$F(x, t)$ satisfies the transformation equation*

$$(2.15) \quad F(x, t) = w(h, k, t) \exp\{\pi(-tz + (D - T)/Tz)/12k\} G(y, t, k, h) \qquad \textit{if } 2 \mid k^*,$$

$$F(x,t)=2^{\frac{1}{2}}W(h,k,t)\exp\{\pi(-tz-(2T+D)/2Tz)/12k\}H(Y,t,k,h) \quad \text{if } 2\mid k,\ 2\nmid k^*, \tag{2.16}$$

$$F(x,t)=z(h,k,t)\exp\{\pi(-tz+(T-D)/2Tz)/12k\}J(X,t,k,h) \quad \text{if } 2\nmid k. \tag{2.17}$$

Here $(k,t+1)=D$, $k=Dk^*$, $t+1=DT$. $w(h,k,t)$, $W(h,k,t)$, $z(h,k,t)$ are given by (2.6), (2.9), (2.12) respectively. $G(y,t,k,h)$, $H(Y,t,k,h)$, $J(X,t,k,h)$ are given by (2.8), (2.11), (2.14) respectively.

3. The exponential sums. In this section we shall be concerned with estimating the magnitude of certain exponential sums involving $w(h,k,t)$, $W(h,k,t)$ and $z(h,k,t)$. Our argument will rest on the following results concerning $\omega(h,k)$ and $\chi(h,k)$ which were established in Section 3 of [1].

If $4\mid k$ then

$$\omega(h,k)=\exp\{2\pi i(\phi(uh+vh')/Gk-6\phi(c-2)/G)\}, \tag{3.1}$$

where if $3\mid k$ then $f=1$, $G=48$, while if $3\nmid k$ then $f=3$, $G=16$. ϕ satisfies the congruence $f\phi\equiv 1\pmod{Gk}$, while $hh'\equiv -1\pmod{Gk}$. $c=1,3$ according as $h\equiv 1,3\pmod 4$, and $u=2k^2-6k-2$, $v=k^2+2$.

If $k\equiv 2\pmod 4$ then

$$\begin{aligned}\omega(h,k)=\exp\{2\pi i(&\lambda(-2h+2h')/Jk\\ &+\Gamma(-12d(k+k/2)+4d+6k-3k^2)/e)\}.\end{aligned} \tag{3.2}$$

Here if $3\mid k$ then $e=16$, $J=3$, while if $3\nmid k$ then $e=48$, $J=1$. λ and Γ are defined by the congruences $e\lambda\equiv 1\pmod{Jk/2}$, $J\Gamma k/2\equiv 1\pmod e$, while $hh'\equiv -1\pmod{Jk}$. $d=\pm 3$ according as $h\equiv \pm 3\pmod 4$. Since we easily verify that $-36(k+k/2)+12\equiv 36(k+k/2)-12\pmod e$, the apparent dependence of $\Gamma(-12d(k+k/2)+4d+6k-3k^2)/e$ on h is illusory.

If $2\nmid k$ then

$$\chi(h,k)=\exp\{2\pi i(\Phi(-2h-h')/gk+\gamma(12k-12k^2+6k\alpha-6)/F)\}. \tag{3.3}$$

Here if $3\mid k$ then $F=16$, $g=3$, while if $3\nmid k$ then $F=48$, $g=1$. Φ and γ are defined by the congruences $F\Phi\equiv 1\pmod{gk}$, $gk\gamma\equiv 1\pmod F$, while $hh'\equiv -1\pmod{gk}$. $\alpha=1,3$ according as $k\equiv 1,3\pmod 4$.

From these statements similar results concerning $w(h,k,t)$, $W(h,k,t)$ and $z(h,k,t)$ can be derived. First we introduce several new symbols. $C(t,k)$ will be used to denote *any* parameter whose value depends on t and k but not on h. $C(t,k,h)$ will denote a parameter dependent on t, k and h but with the property that $C(t,k,h_1)=C(t,k,h_2)$ if $h_1\equiv h_2\pmod 4$. When used in different expressions $C(t,k)$ and $C(t,k,h)$ need not have the same value. The Roman numeral I will symbolize the condition that either $3 \mid k$ *and* $3 \mid k^*$, or $3 \nmid k$. The Roman numeral II will symbolize the condition that $3 \mid k$ *and* $3 \nmid k^*$.

For $w(h,k,t)$ we must consider three cases. If $4 \mid k^*$ then from (3.1)

$$(3.4) \qquad \omega(Th,k^*) = \exp\{2\pi i(\phi^*(u^*Th + v^*h^*)/G^*k^* + C(t,k,h))\}.$$

Here if I then $f^* = f$, $G^* = G$, $\phi^* = \phi$; while if II then $f^* = 3f$, $3G^* = G$, $\phi^* \equiv a_1\phi \pmod{G^*k^*}$ where $3a_1 \equiv 1 \pmod{G^*k^*}$. $Thh^* \equiv -1 \pmod{G^*k^*}$ so that if $TT' \equiv 1 \pmod{G^*k^*}$ we have $h^* \equiv T'h' \pmod{G^*k^*}$.

$$u^* = 2k^{*2} - 6k^* - 2, \qquad v^* = k^{*2} + 2.$$

From (2.6), (3.1), (3.4) we conclude that if $4 \mid k^*$, then

$$(3.5) \qquad w(h,k,t) = \exp\{2\pi i(\phi(uh + vh')/Gk + C(t,k,h))\}.$$

If I then

$$u = 2k^2 - 6k - (t+1)(2k^{*2} - 6k^*) + 2t, \; v = k^2 + 2 - DT'(2 + k^{*2}).$$

If II then

$$u = 2k^2 - 6k - 2 - 3a_1(t+1)(2k^{*2} - 6k^* - 2),$$
$$v = k^2 + 2 - 3a_1DT'(2 + k^{*2}).$$

ϕ and G are defined in (3.1).

If $4 \mid k$ and $k^* \equiv 2 \pmod 4$ then from (3.2)

$$(3.6) \qquad \omega(Th,k^*) = \exp\{2\pi i(\lambda^*(-2Th + 2h^*)/J^*k^* + C(t,k))\}$$

Here if I then $e^* = 16f$, $16J^* = G$, $\lambda^* \equiv a_2\phi \pmod{J^*k^*/2}$ where $16a_2 \equiv 1 \pmod{J^*k^*/2}$, while if II then $e^* = 48f$, $48J^* = G$, $\lambda^* \equiv a_3\phi \pmod{J^*k^*/2}$ where we define a_3 by $48a_3 \equiv 1 \pmod{J^*k^*/2}$. $Thh^* \equiv -1 \pmod{J^*k^*/2}$ so that if $TT' \equiv 1 \pmod{J^*k^*}$ we have $h^* \equiv T'h' \pmod{J^*k^*/2}$. From (2.6), (3.1), (3.6) we conclude that if $4 \mid k$ and $k^* \equiv 2 \pmod 4$ then

$$(3.7) \qquad w(h,k,t) = \exp\{2\pi i(\phi(uh + vh')/Gk + C(t,k,h))\}.$$

If I then $u = 2k^2 - 6k - 2 + 32a_2(t+1)$, $v = k^2 + 2 - 32a_2DT'$. If II then $u = 2k^2 - 6k - 2 + 96a_3(t+1)$, $v = k^2 + 2 - 96a_3DT'$. ϕ and G are defined in (3.1).

If $k \equiv k^* \equiv 2 \pmod 4$ then from (3.2)

$$(3.8) \qquad \omega(Th,k^*) = \exp\{2\pi i(\lambda^*(-2Th + 2h^*)/J^*k^* + C(t,k))\}.$$

Here if I then $e^*=e$, $J^*=J$, $\lambda^*=\lambda$; while if II then $e^*=3e$, $3J^*=J$, $\lambda^*\equiv a_4\lambda \pmod{J^*k^*/2}$ where $3a_4\equiv 1 \pmod{J^*k^*/2}$. As before, $h^*\equiv T'h'$ $\pmod{J^*k^*/2}$. From (2.6), (3.2), (3.8) we conclude that if $k\equiv k^*\equiv 2$ $\pmod 4$ then

$$(3.9) \qquad w(h,k,t)=\exp\{2\pi i(\lambda(uh+vh')/Jk+C(t,k))\}.$$

If I then $u=2t$, $v=2-2DT'$. If II then $u=-2+6a_4(t+1)$, $v=2-6a_4DT'$. λ and J are defined in (3.2).

Turning to $W(h,k,t)$ we must consider two cases. If $4\mid k$ and $2\nmid k^*$ then from (3.3)

$$(3.10) \qquad \chi(Th,k^*)=\exp\{2\pi i(\Phi^*(-2Th-h^*)/g^*k^*+C(t,k))\}.$$

If I then $F^*=16f$, $16g^*=G$, $\Phi^*\equiv a_5\phi \pmod{g^*k^*}$ where $16a_5\equiv 1 \pmod{g^*k^*}$; while if II then $F^*=48f$, $48g^*=G$, $\Phi^*\equiv a_6\phi \pmod{g^*k^*}$ where $48a_6\equiv 1$ $\pmod{g^*k^*}$. $Thh^*\equiv -1 \pmod{g^*k^*}$ so that if $TT'\equiv 1 \pmod{g^*k^*}$ we have $h^*\equiv T'h' \pmod{g^*k^*}$. From (2.9), (3.1), (3.10) we have if $4\mid k$ and $2\nmid k^*$

$$(3.11) \qquad W(h,k,t)=\exp\{2\pi i(\phi(uh+vh')/Gk+C(t,k,h))\}.$$

If I then $u=2k^2-6k-2+32a_5(t+1)$, $v=k^2+2+16a_5DT'$. If II then $u=2k^2-6k-2+96a_6(t+1)$, $v=k^2+2+48a_6DT'$. G and ϕ are defined in (3.1).

If $k\equiv 2 \pmod 4$ and $2\nmid k^*$ we have

$$(3.12) \qquad \chi(Th,k^*)=\exp\{2\pi i(\Phi^*(-2Th-h^*)/g^*k^*+C(t,k))\}.$$

Here if I then $F^*=e$, $g^*=J$, $\Phi^*=\lambda$; while if II then $F^*=3e$, $3g^*=J$, $\Phi^*\equiv a_7\lambda \pmod{g^*k^*}$ where $3a_7\equiv 1 \pmod{g^*k^*}$. As before $h^*\equiv T'h' \pmod{g^*k^*}$. From (2.9), (3.2), (3.12) we conclude that if $k\equiv 2 \pmod 4$ and $2\nmid k^*$ then

$$(3.13) \qquad W(h,k,t)=\exp\{2\pi i(\lambda(uh+vh')/Jk+C(t,k))\}.$$

If I then $u=2t$, $v=2+DT'$. If II then $u=-2+6a_7(t+1)$, $v=2+3a_7DT'$. λ and J are defined in (3.2).

If $2\nmid k$ then from (3.3)

$$(3.14) \qquad \chi(Th,k^*)=\exp\{2\pi i(\Phi^*(-2Th-h^*)/g^*k^*+C(t,k))\}.$$

If I then $F^*=F$, $g^*=g$, $\Phi^*=\Phi$; while if II then $F^*=3F$, $3g^*=g$, $\Phi^*\equiv a_7\Phi \pmod{g^*k^*}$ where a_7 is defined as before. $h^*\equiv T'h' \pmod{g^*k^*}$. From (2.12), (3.3), (3.14)

$$(3.15) \qquad z(h,k,t)=\exp\{2\pi i(\Phi(uh+vh')/gk+C(t,k))\}.$$

If I then $u=2t$, $v=-1+DT'$. If II then $u=-2+6a_7(t+1)$, $v=-1+3a_7DT'$. Φ and g are defined in (3.3).

We now state the main result of this section. The proof depends on the identities just established.

THEOREM 2. *If* $(k,t+1)=D$, $h\equiv a(\operatorname{mod} D)$, $(a,D)=1$; $s_1\leqq h'<s_2$ $(\operatorname{mod} k)$, $0\leqq s_1<s_2\leqq k$; $n\geqq t$; *then the sums*

$$A_1=\sum_{h \operatorname{mod} k}{}' w(h,k,t)\exp\{-2\pi i(hn-h'v)/k\} \text{ where } 2\mid k^*;$$

$$A_2=\sum_{h \operatorname{mod} k}{}' W(h,k,t)\exp\{-2\pi i(hn-h'v)/k\} \text{ where } 2\mid k,\ 2\nmid k^*;$$

$$A_3=\sum_{h \operatorname{mod} k}{}' z(h,k,t)\exp\{-2\pi i(hn-H'v)/k\} \text{ where } 2\nmid k;$$

are each subject to the estimate $0(n^{1/3}k^{2/3+\epsilon})$ *uniformly in* a, v, s_1, s_2. $k^*=k/D$, $hh'\equiv-1(\operatorname{mod} k)$, $2hH'\equiv-1(\operatorname{mod} k)$ *and* $\sum'$ *indicates that* h *runs over a reduced residue system modulo* k.

Proof. For A_1 there are two cases to be considered. From (2.6) and (2.3) we see that $w(h,k,t)$ has period k when viewed as a function of h. Therefore, if $4\mid k$ and we change the modulus to Gk in A_1 and select h' so that $hh'\equiv-1(\operatorname{mod} Gk)$ we see that

$$A_1=G^{-1}\sum_{h \operatorname{mod} Gk}{}'\exp\{2\pi i(C(t,k,h)+f(h)/Gk)\}$$

where we have used (3.5), (3.7) and where

$$f(h)=(\phi u-Gn)h+(\phi v+Gv)h'.$$

If we split the sum over h into two parts according as $h\equiv1(\operatorname{mod}4)$ or $h\equiv3(\operatorname{mod}4)$ then in each part $C(t,k,h)$ is a constant, and we have $A_1=A_{11}+A_{13}$ where $A_{1d}=c_d\sum_{h \operatorname{mod} Gk}{}'\exp\{2\pi if(h)/Gk\}$ and $h\equiv d(\operatorname{mod}4)$, $|c_d|=G^{-1}$.

If we define the function $m(s)$ for all integers s by

$$m(s)=\begin{cases}1 \text{ if } s_1\leqq s<s_2(\operatorname{mod} k),\\ 0 \text{ otherwise,}\end{cases}$$

then $m(s)$ has period k and from the theory of finite Fourier series we have $m(s)=\sum_{j=0}^{k-1}\lambda_j\exp\{2\pi isj/k\}$ where $\lambda_j=k^{-1}\sum_{s=0}^{k-1}m(s)\exp\{-2\pi isj/k\}$. We easily verify (see Section 2 in [3]) that $\sum_{j=0}^{k-1}\lambda_j=0(\log k)$, from which it follows that

$$\sum_{j=0}^{k-1} \lambda_j = 0(k^{\epsilon}) \text{ for any } \epsilon > 0. \tag{3.16}$$

We can now write

$$\begin{aligned} A_{1d} &= c_d \sum_{h \bmod Gk}{}' m(h') \exp\{2\pi i f(h)/Gk\} \\ &= c_d \sum_{j=0}^{k-1} \lambda_j \sum_{h \bmod Gk}{}' \exp\{2\pi i((\phi u - Gn)h + (\phi v + G\nu + Gj)h')/Gk\} \end{aligned} \tag{3.17}$$

where $h \equiv a (\bmod D)$, $h \equiv d (\bmod 4)$ and the restriction $s_1 \leqq h' < s_2 (\bmod k)$ has been dropped. If $4 \mid D$ and $a \not\equiv d (\bmod 4)$ then $A_{1d} = 0$. Otherwise the two congruence conditions on h can be replaced by $h \equiv d^* (\bmod [4, D])$ where $[4, D]$ is the lowest common multiple of 4 and D. In this case the inner sum in A_{1d} is a generalized Kloosterman sum and by a theorem of Salié [6] is subject to the estimate $0((\phi u - Gn, Gk)^{1/3} (Gk)^{2/3+\epsilon})$. Therefore, from (3.16), (3.17) and the fact that $A_1 = A_{11} + A_{13}$,

$$A_1 = 0((\phi u - Gn, Gk)^{1/3} k^{2/3+\epsilon})$$

where the multiplicative constant implied by 0 is independent of k, n, a, ν, s_1, s_2.

Since $(f, Gk) = 1$ (see (3.1)) we have $(\phi u - Gn, Gk) = (f\phi u - fGn, Gk)$. But $fG = 48$ and $f\phi u = u + mGk$ where m is an integer. Therefore, $(\phi u - Gn, Gk) = (u - 48n, Gk) \leqq DG(u - 48n, k^*)$. From (3.5), (3.7) (recall that $3a_1 = 1 + mk^*$, $32a_2 = 2 + mk^*$, $96a_3 = 2 + mk^*$) we see that $(u - 48n, k^*) = (2t - 48n, k^*) = 0(n)$. Therefore, $A_1 = 0(n^{1/3} k^{2/3+\epsilon})$ as was to be proved.

If $4 \nmid k$ we change the modulus to Jk in A_1 and select h' so that $hh' \equiv -1 (\bmod Jk)$. From (3.9)

$$A_1 = J^{-1} \sum_{h \bmod Jk} \exp\{2\pi i(C(t, k) + g(h)/Jk)\}$$

where $g(h) = (\lambda u - Jn)h + (\lambda v + J\nu)h'$. Since $C(t, k)$ is independent of h we have, defining $m(s)$ as before

$$A_1 = c \sum_{j=0}^{k-1} \lambda_j \sum_{h \bmod Jk}{}' \exp\{2\pi i(g(h) + jJh')/Jk\}$$

where $|c| = J^{-1}$, $h \equiv a (\bmod D)$ and the restriction $s_1 \leqq h' < s_2 (\bmod k)$ has been removed. The inner sum is a generalized Kloosterman sum, and as before $A_1 = 0((\lambda u - Jn, Jk)^{1/3} k^{2/3+\epsilon})$.

Now $(\lambda u - Jn, Jk) \leqq (2e\lambda u - 2eJn, Jk)$. But $eJ = 48$ and $2e\lambda u = 2u + mJk$ (see (3.2)). Therefore, $(\lambda u - Jn, Jk) \leqq DJ(2u - 96n, k^*)$. From (3.9) (recall that $6a_4 = 2 + mk^*$) we see that

$$(2u-96n,k^*)=(4t-96n,k^*)=0(n).$$

We conclude that $A_1=0(n^{1/3}k^{2/3+\epsilon})$.

The treatment of A_2 and A_3 is similar to that of A_1. We shall not give the details.

4. A convergent series for $q(n,t)$. In what follows we assume that $n\geqq t$. By Cauchy's formula we have

$$q(n,t)=\frac{1}{2\pi i}\int_C x^{-n-1}F(x,t)\,dx=\sum_{h,k}{}'\frac{1}{2\pi i}\int_{\xi_{hk}}x^{-n-1}F(x,t)\,dx.$$

Here $0\leqq h<k\leqq N$, $(h,k)=1$, and ξ_{hk} are the Farey arcs of order N of C, the circle $|x|=\exp\{-2\pi N^{-2}\}$. If on ξ_{hk} we let $x=\exp\{2\pi ih/k-2\pi z/k\}$ where $z=wk$, $w=N^{-2}-i\phi$ we obtain

$$q(n,t)=\sum_{h,k}{}'\exp\{-2\pi inh/k\}\int_{-\alpha}^{\beta}F(\exp\{2\pi ih/k-2\pi z/k\},\,t)\exp\{2\pi nw\}\,d\phi.$$

Here $\alpha=1/k(k+k_1)$ and $\beta=1/k(k+k_2)$ where k_1, k, k_2 are the denominators of consecutive terms in the Farey series of order N.

We now split the sum over k into $\tau(t+1)$ parts, where as usual $\tau(m)$ is the number of positive divisors of m. We denote these parts as $S_1,\cdots,S_D,\cdots,S_{t+1}$ where in S_D we include all k such that $(k,t+1)=D$. If we split S_D into three parts S_{D1}, S_{D2}, S_{D3} according as $2\mid k^*$; $2\mid k$ and $2\nmid k^*$, $2\nmid k$ respectively then

$$q(n,t)=\sum_{D\mid t+1}\sum_{j=1}^{3}S_{Dj}.$$

For a fixed value of D at least one, and perhaps two, of the S_{Dj} is empty. Thus, a sufficient condition that S_{D3} be empty is that $2\mid D$, while S_{D2} is empty if $2\nmid D$. S_{D1} and S_{D2} are both empty if $t+1$ is even and D is odd. S_{D1} and S_{D3} are both empty if $t+1=(2m+1)2^j$, $m\geqq 0$, $j\geqq 2$ and $2\mid D$, $2^j\nmid D$.

Considering first S_{D1} we have $S_{D1}=\sum\limits_{b \bmod D}{}'\sum\limits_{d=1}^{D}S_{D1}(b,d)$ where $S_{D1}(b,d)$ indicates that we are including only those h and k such that $h'\equiv b \pmod{D}$ where $hh'\equiv -1 \pmod{D}$, and $m\equiv d \pmod{D}$ where $1-TT'=mk^*$. (Recall that T' is a fixed solution of $TT'\equiv 1 \pmod{k^*}$.) Either $S_{D1}(b,d)=0$ or we have from (2.15), (2.7), (2.8)

$$S_{D1}(b,d) = \sum_{h,k}' w(h,k,t)\exp\{-2\pi inh/k\}$$
$$\cdot \int_{-\alpha}^{\beta} \sum_{\nu=0}^{\infty} c(\nu,b,d,t,k)\exp\{2\pi ih'T'\nu/k\}$$
$$\cdot \exp\{-(\pi/k^2wT)(2\nu-(D-T)/12)+\pi w(2n-t/12)\}d\phi$$

where $1 \leqq k \leqq N$, $k = Dk^*$, $2 \mid k^*$ and $h \equiv a (\operatorname{mod} D)$ (if $ab \equiv -1 (\operatorname{mod} D)$). Note that $c(\nu,b,d,t,k)$ depends *only* on ν here. Splitting the sum over ν into two parts $Q_{D1}(b,d)$ and $R_{D1}(b,d)$ according as $\nu < (D-T)/24$ or $\nu \geqq (D-T)/24$ we find, using Rademacher's argument [5] and making use of the estimate for A_1 given in Theorem 2, that

$$R_{D1}(b,d) = 0(n^{1/3}N^{-1/3+\epsilon}\exp\{2\pi nN^{-2}\}).$$

Since $Q_{D1}(b,d)$ contains all ν such that $\nu < (D-T)/24$ we see that if $D \leqq T$ then the sum over ν is vacuous and $Q_{D1}(b,d) = 0$. Since $DT = t+1$ we see that this will occur if and only if $D \leqq (t+1)^{1/2}$. We therefore restrict our attention to those D such that $D > (t+1)^{1/2}$.

Proceeding as in [5] we find that

$$Q_{D1}(b,d) = 2\pi \sum_{k}\sum_{\nu} c(\nu,b,d,t,k)A(k,t,n,\nu,a,T')L'(k,t,n,\nu)$$
$$+ 0(n^{1/3}N^{-1/3+\epsilon}\exp\{2\pi nN^{-2}\})$$

where $1 \leqq k \leqq N$, $k = Dk^*$, $2 \mid k^*$, $t+1 = DT$, $\nu < (D-T)/24$, $1 - TT' = mk^*$, $m \equiv d (\operatorname{mod} D)$.

$$A(k,t,n,\nu,a,T') = \sum_{h \bmod k}' w(h,k,t)\exp\{-2\pi i(nh - T'\nu h')/k\}$$

where $h \equiv a (\operatorname{mod} D)$.

$$L'(k,t,n,\nu) = k^{-1}\{(D-T-24\nu)/(24nT-tT)\}^{1/2}$$
$$\cdot I_1\{\pi(24n-t)^{1/2}(D-T-24\nu)^{1/2}/6kT^{1/2}\}$$

where $I_1(z)$ is the Bessel function of order one.

Since $T < D$, we see from (2.8) that if $\nu < (D-T)/24$ then

$$c(\nu,b,d,t,k) = \exp\{2\pi i\nu bd/TD\}q(\nu/T)$$
$$= \exp\{2\pi i\nu h'(1-TT')/kT\}q(\nu/T)$$

if $T \mid \nu$, and $c(\nu,b,d,t,k) = 0$ otherwise. Therefore, if we let $\nu = Tm$ and write

$$J = (D-T)/24T, \qquad r = t/24 \tag{4.1}$$

we have, recalling that $S_{D1}(b, d) = Q_{D1}(b, d) + R_{D1}(b, d)$, and summing over b and d,

$$S_{D1} = 2\pi \sum_{k=1}^{N} \sum_{m<J} q(m) A(k, t, n, m) L(k, t, n, m) \qquad (4.2)$$
$$+ 0(n^{1/3} N^{-1/3+\epsilon} \exp\{2\pi n N^{-2}\}$$

where $(k, t+1) = D$, $k = Dk^*$, $2 \mid k^*$.

$$A(k, t, n, m) = \sum_{h \bmod k}' w(h, k, t) \exp\{-2\pi i(nh - mh')/k\}, \qquad (4.3)$$

$$L(k, t, n, m) = k^{-1}\{(J - m)/(n - r)\}^{1/2} I_1\{4\pi(n - r)^{1/2}(J - m)^{1/2}/k\}. \qquad (4.4)$$

In (4.2) and the remainder of this section the multiplicative constant implied by the 0 notation depends at most on t.

Turning now to S_{D2} we have $S_{D2} = \sum_{b \bmod D}' \sum_{d=1}^{D} S_{D2}(b, d)$ where $S_{D2}(b, d)$ includes those terms of S_{D2} for which $h' \equiv b \pmod D$ and $m \equiv d \pmod D$ where $1 - 2TT^* = mk^*$. (Recall that T^* is a fixed solution of $2TT^* \equiv 1 \pmod{k^*}$.) Either $S_{D2}(b, d) = 0$ or from (2.16), (2.10), (2.11)

$$S_{D2}(b, d) = 2^{1/2} \sum_{h,k}' W(h, k, t) \exp\{-2\pi i n h/k\}$$
$$\cdot \int_{-\alpha}^{\beta} \sum_{\nu=0}^{\infty} d(\nu, b, d, t, k) \exp\{2\pi i h' T^* \nu/k\}$$
$$\cdot \exp\{-(\pi/2k^2 wT)(2\nu + (2T + D)/12) + \pi w(2n - t/12)\} d\phi$$

where $1 \leqq k \leqq N$, $k = Dk^*$, $2 \mid k$, $2 \nmid k^*$ and $h \equiv a \pmod D$ $(ab \equiv -1 \pmod D)$. Note that $d(\nu, b, d, t, k)$ depends *only* on ν here. Since $2\nu + (2T + D)/12 > 0$ we find, employing the estimate of A_2 given in Theorem 2, that $S_{D2}(b, d) = 0(n^{1/3} N^{-1/3+\epsilon} \exp\{2\pi n N^{-2}\})$. Summing over b and d we have

$$S_{D2} = 0(n^{1/3} N^{-1/3+\epsilon} \exp\{2\pi n N^{-2}\}). \qquad (4.5)$$

Writing $S_{D3} = \sum_{b \bmod D}' \sum_{d=1}^{D} S_{D3}(b, d)$, where $S_{D3}(b, d)$ includes those terms of S_{D3} for which $H' \equiv b \pmod D$ and $m \equiv d \pmod D$ where $1 - TT' = mk^*$, we have either $S_{D3}(b, d) = 0$ or from (2.17), (2.13), (2.14)

$$S_{D3}(b, d) = \sum_{h,k}' z(h, k, t) \exp\{-2\pi i n h/k\}$$
$$\cdot \int_{-\alpha}^{\beta} \sum_{\nu=0}^{\infty} e(\nu, b, d, t, k) \exp\{2\pi i H' T' \nu/k\}$$
$$\cdot \exp\{-(\pi/k^2 wT)(\nu - (T - D)/24) + \pi w(2n - t/12)\} d\phi$$

where $1 \leqq k \leqq N$, $k = Dk^*$, $2 \nmid k$ and $h \equiv a (\operatorname{mod} D)$ (where $2ab \equiv -1 (\operatorname{mod} D)$). Note that $e(\nu, b, d. t, k)$ depends *only* on ν here. If we split the sum over ν into two parts $Q_{D3}(b, d)$ and $R_{D3}(b, d)$ according as $\nu < (T-D)/24$ or $\nu \geqq (T-D)/24$ we find, employing the estimate of A_3 given in Theorem 2, that

$$R_{D3}(b, d) = 0(n^{1/3}N^{-1/3+\epsilon}\exp\{2\pi nN^{-2}\}).$$

Since $Q_{D3}(b, d)$ contains all ν such that $\nu < (T-D)/24$ we see that if $D \geqq T$ then $Q_{D3}(b, d) = 0$. This will occur if and only if $D \geqq (t+1)^{1/2}$. Therefore, we restrict our attention to $D < (t+1)^{1/2}$.

Following Rademacher's method [5] and employing Theorem 2 we have

$$Q_{D3}(b, d) = 2\pi \sum_{k} \sum_{\nu} e(\nu, b, d, t, k) B(k, t, n, \nu, a, T') M'(k, t, n, \nu) \\ + 0(n^{1/3}N^{-1/3+\epsilon}\exp\{2\pi nN^{-2}\})$$

where $1 \leqq k \leqq N$, $k = Dk^*$, $2 \nmid k$, $t + 1 = DT$, $\nu < (T-D)/24$, $1 - TT' = mk^*$, $m \equiv d (\operatorname{mod} D)$.

$$B(k, t, n, \nu, a, T') = \sum_{h \operatorname{mod} k}' z(h, k, t) \exp\{-2\pi i(nh - T'\nu H')/k\}$$

where $h \equiv a (\operatorname{mod} D)$.

$$M'(k, t, n, \nu) = k^{-1}\{(T - D - 24\nu)/(48nT - 2tT)\}^{1/2} \\ \cdot I_1\{\pi(48n - 2t)^{1/2}(T - D - 24\nu)^{1/2}/12kT^{1/2}\}.$$

Since $D < T$ we see from (2.4) that if $\nu < (T-D)/24$ then $e(\nu, b, d, t, k) = q(\nu/D)$ if $D \mid \nu$ and $e(\nu, b, d, t, k) = 0$ otherwise. Therefore, if we let $\nu = Dm$, and write

$$K = (T - D)/24D, \quad r = t/24, \quad \mu = (D/2T)^{1/2} \tag{4.6}$$

we have, recalling that $S_{D3}(b, d) = Q_{D3}(b, d) + R_{D3}(b, d)$ and summing over b and d

$$S_{D3} = 2\pi \sum_{k-1}^{N} \sum_{m<K} q(m) B(k, t, n, m) M(k, t, n, m) \\ + 0(n^{1/3}N^{-1/3+\epsilon}\exp\{2\pi nN^{-2}\} \tag{4.7}$$

where $(k, t+1) = D$, $k = Dk^*$, $2 \nmid k$.

$$B(k, t, n, m) = \sum_{h \operatorname{mod} k}' z(h, k, t) \exp\{-2\pi i(nh - DT'mH')/k\}, \tag{4.8}$$

$$M(k, t, n, m) \\ = \mu k^{-1}\{(K - m)/(n - r)\}^{1/2} I_1\{4\mu\pi(n - r)^{1/2}(K - m)^{1/2}/k\}. \tag{4.9}$$

We note that B does *not* depend on T' since $DT'mH'/k = T'mH'/k^*$ and T' is determined modulo k^* by k and t.

Summing over D, and then letting $N \to \infty$ in (4.2), (4.5), (4.7) we have

THEOREM 3. *$q(n,t)$, the number of partitions of n into odd summands where no summand is used more than t times, is given by the convergent series*

$$q(n,t) = 2\pi\Big(\sum_{D}\sum_{k>0}\sum_{m<J} q(m)A(k,t,n,m)L(k,t,n,m) + \sum_{D}\sum_{k>0}\sum_{m<K} q(m)B(k,t,n,m)M(k,t,n,m)\Big). \tag{4.10}$$

Here $D \mid (t+1)$ and in the first summation $D > (t+1)^{1/2}$, in the second summation $D < (t+1)^{1/2}$. $(k, t+1) = D$, $k = Dk^*$ and in the first summation $2 \mid k^*$, in the second summation $2 \nmid k$. J, A, L, K, B, M are given by (4.1), (4.3), (4.4), (4.6), (4.8), (4.9) respectively. $q(m)$ is the number of partitions of m into odd parts.

5. Some special cases. If $t = 2^j - 1$, $j \geqq 1$ then only $D = 2^j$ in the first summation of (4.10) and $D = 1$ in the second summation yield acceptable values of k. In this case $D = t+1$, $T = 1$, $J = r = t/24$ in the first summation and $D = 1$, $T = t+1$, $K = r$ in the second. Thus, we have

COROLLARY 3.1. *If $t = 2^j - 1$, $j \geqq 1$ then*

$$q(n,t) = 2\pi\Big(\sum_{k}\sum_{m<r} q(m)A(k,t,n,m)L(k,t,n,m) + \sum_{2\nmid k}\sum_{m<r} q(m)B(k,t,n,m)M(k,t,n,m)\Big)$$

where $2^{j+1} \mid k$ in the first summation.

In particular, if $t = 1$, then $r = 1/24$ and $q(n,1) = Q(n)$ where $Q(n)$ is the number of partitions of n into odd and unequal parts. In this case we have

COROLLARY 3.2. *The number of partitions of a positive integer n into odd and unequal parts is given by*

$$Q(n) = 2\pi\Big(\sum_{4\mid k} A(k,1,n,0)L(k,1,n,0) + \sum_{2\nmid k} B(k,1,n,0)M(k,1,n,0)\Big)$$

This result is easily seen to agree with Theorem 5 of [2].

If $t = p^j - 1$ where p is an odd prime and $j = 1$ or $j = 2$, then $D = p^j$, $T = 1$, $J = r = t/24$ in the first summation of (4.10), and $D = 1$, $T = p^j$, $K = r$ in the second summation. We then have

Corollary 3.3. *If $t = p^j - 1$, p an odd prime, $j = 1$ or $j = 2$, then*

$$q(n,t) = 2\pi\Big(\sum_{k}\sum_{m<r} q(m)A(k,t,n,m)L(k,t,n,m)$$
$$+ \sum_{k}\sum_{m<r} q(m)B(k,t,n,m)L(k,t,n,m)\Big).$$

In the first summation $2p^j \mid k$, in the second $(2p,k) = 1$.

We remark further that if $p^j < 73$ then $m < 3$ here and $q(m) = 1$.

6. Asymptotic formulae. If $k = 1$ then $D = 1$, $T = t + 1$, $K = t/24$. Writing

$$r = t/24, \tag{6.1}$$

$$G(m) = (r - m)^{1/2}, \tag{6.2}$$

$$E = 4\pi(n - r)^{1/2}, \tag{6.3}$$

$$s = (2t + 2)^{-1/2}, \tag{6.4}$$

$$W = \sum_{m<r} q(m)G(m)I_1\{sEG(m)\}, \tag{6.5}$$

we have from (4.10)

$$q(n,t) = 2\pi sW(n - r)^{-1/2}\{1 + S_1 + S_2\}. \tag{6.6}$$

Here

$$S_1 = \sum_{D}\sum_{k>1}\sum_{m<J} q(m)A(k,t,n,m)(J - m)^{1/2}(ksW)^{-1}I_1\{E(J - m)^{1/2}/k\} \tag{6.7}$$

where $D \mid (t+1)$, $D > (t+1)^{1/2}$, $(k, t+1) = D$, $k = Dk^*$, $2 \mid k^*$;

$$S_2 = \sum_{D}\sum_{k>1}\sum_{m<K} q(m)B(k,t,n,m)(K - m)^{1/2}\mu(ksW)^{-1}I_1\{\mu E(K - m)^{1/2}/k\} \tag{6.8}$$

where $D \mid (t+1)$, $D < (t+1)^{1/2}$, $(k, t+1) = D$, $k = Dk^*$, $2 \nmid k$.

We now proceed to obtain estimates of the magnitude of S_1 and S_2 First we have from the theory of Bessel functions: if x is a positive real variable then

(6.9) $I_1(x)$ is a positive, monotonic increasing function of x,

(6.10) $I_1(x) = 0(x)$ if $x < 1$,

(6.11) $I_1(x) = e^x(2\pi x)^{-1/2}(1 + 0(x^{-1}))$ if $x > 1$.

We shall also require

Lemma 1. *If W is given by (6.5), then for "large" n*

$$W = r^{1/2} I_1\{sEr^{1/2}\}(1 + 0(\exp\{-cn^{1/2}\})) \tag{6.12}$$

where $c > 0$.

Proof. If $r \leqq 1$ there is nothing to prove. We assume therefore that $t > 24$. From (6.5) we have

$$\begin{aligned} W = r^{1/2} I_1\{sEr^{1/2}\}(1 + \sum_{0<m<r} q(m)(1 - m/r)^{1/2} \\ \cdot I_1\{sEG(m)\}/I_1\{sEr^{1/2}\}). \end{aligned} \tag{6.13}$$

For large n it follows from (6.9), (6.11), (6.2), (6.3) that

$$\begin{aligned} I_1\{sEG(m)\}/I_1\{sEr^{1/2}\} &= 0(\exp\{sE(G(1) - r^{1/2})\}) \\ &= 0(\exp\{-4\pi sn^{1/2}(1 - r/n)^{1/2}(r^{1/2} - (r-1)^{1/2})\}). \end{aligned} \tag{6.14}$$

Since $q(m)(1 - m/r)^{1/2}$ is positive and bounded for $0 < m < r$, (6.12) frollows from (6.13) and (6.14).

From (4.3) and Theorem 2 we see that $A(k, t, n, m) = 0(n^{1/3}k^{2/3+\epsilon})$. Also, from (4.1) we have $J \leqq r$. Therefore, from (6.7), (6.9), (6.12) and the boundedness of $q(m)$ for $m \leqq r$ we have for large n

$$S_1 = 0(\sum_{k=2U}^{\infty} n^{1/3}k^{-1/3+\epsilon} I_1\{Er^{1/2}/k\}/I_1\{sEr^{1/2}\}) \tag{6.15}$$

where $U > (t + 1)^{1/2}$.

We now split the sum in (6.15) into two parts according as $k < Er^{1/2} = R$ or $k > R$. Using (6.11) in the first case and (6.10) in the second we have for large n

$$\begin{aligned} S_1 = 0(\sum_{k=2U}^{R} n^{1/3}k^{1/6+\epsilon} \exp\{-Er^{1/2}(s - 1/k)\}) \\ + 0(\sum_{k>R} n^{1/3}k^{-4/3+\epsilon}E^{3/2} \exp\{-Esr^{1/2}\}). \end{aligned} \tag{6.16}$$

Since $k \geqq 2U > 2(t+1)^{1/2}$ we see from (6.4) that in S_1

$$s > s - 1/k > s - s/2^{1/2} = g > 0. \tag{6.17}$$

Also, from (6.3) we have for large n

$$2\pi n^{1/2} < E = 4\pi n^{1/2}(1 - r/n)^{1/2} < 4\pi n^{1/2}. \tag{6.18}$$

From (6.16), (6.17), (6.18) we conclude that

$$\begin{aligned} S_1 = 0(n^{1/3}n^{7/12+\epsilon} \exp\{-2\pi g r^{1/2} n^{1/2}\}) \\ + 0(n^{1/3}n^{-1/6+\epsilon}n^{3/4} \exp\{-2\pi g r^{1/2} n^{1/2}\}), \end{aligned}$$

or $S_1 = 0(n \exp\{-cn^{1/2}\})$. Since $n = 0(\exp\{\epsilon n^{1/2}\})$ for any $\epsilon > 0$, we have for large n

$$S_1 = 0(\exp\{-c_1 n^{1/2}\}) \text{ where } c_1 > 0. \tag{6.19}$$

Turning now to S_2 we have from (4.8) and Theorem 2,

$$B(k, t, n, m) = 0(n^{1/3}k^{2/3+\epsilon}).$$

From (4.6) we have $K \leqq r$, $0 < \mu < 1$, and $0 < \mu K^{1/2} < (48)^{-1/2}$. Therefore, from (6.8), (6.9), (6.12) and the boundedness of $q(m)$ for $m \leqq r$ we have for large n

$$S_2 = 0\Big(\sum_{k=3}^{\infty} n^{1/3}k^{-1/3+\epsilon} I_1\{E\phi/k\}/I_1\{sEr^{1/2}\}\Big) \tag{6.20}$$

where

$$\phi = (48)^{-1/2}. \tag{6.21}$$

Splitting the sum in (6.20) into two parts according as $k < E\phi = V$ or $k > V$, and again employing (6.11) and (6.10) we have

$$\begin{aligned} S_2 = 0\Big(\sum_{k=3}^{V} n^{1/3}k^{1/6+\epsilon} \exp\{-E(sr^{1/2} - \phi/k)\}\Big) \\ + 0\Big(\sum_{k>V} n^{1/3}k^{-4/3+\epsilon} E^{3/2} \exp\{-Esr^{1/2}\}\Big). \end{aligned} \tag{6.22}$$

Since $sr^{1/2} = \phi(t/(t+1))^{1/2} \geqq \phi/2^{1/2}$ and $-\phi/k \geqq -\phi/3$ in S_2, we have

$$sr^{1/2} - \phi/k \geqq \phi(2^{-1/2} - 1/3) = G > 0 \tag{6.23}$$

where ϕ is given in (6.21).

From (6.22), (6.23), (6.17), (6.18) we have

$$\begin{aligned} S_2 = 0(n^{1/3}n^{7/12+\epsilon} \exp\{-2\pi G n^{1/2}\}) \\ + 0(n^{1/3}n^{-1/6+\epsilon}n^{3/4} \exp\{-2\pi g r^{1/2} n^{1/2}\}) \end{aligned}$$

from which it follows that

$$S_2 = 0(\exp\{-c_2 n^{1/2}\}) \text{ where } c_2 > 0. \tag{6.24}$$

From (6.6), (6.19), (6.24) we obtain

THEOREM 4. *As* $n \to \infty$

$$q(n, t) = 2\pi s W(n - r)^{-1/2}(1 + 0(\exp\{-Cn^{1/2}\})) \tag{6.25}$$

where $C > 0$ *and* r, s, W *are given by* (6.1), (6.4), (6.5) *respectively.*

From (6.25) and (6.12) we have

COROLLARY 4.1. *As* $n \to \infty$

(6.26) $$q(n,t) = 2\pi s r^{1/2}(n-r)^{-1/2} I_1\{sEr^{1/2}\}(1 + 0(\exp\{-Cn^{1/2}\}))$$

where $C > 0$.

Finally, from (6.26) and (6.11) we have, after simplification,

COROLLARY 4.2. *As* $n \to \infty$

(6.27) $$q(n,t) = (s/2)^{1/2} r^{1/4}(n-r)^{-3/4} \exp\{sEr^{1/2}\}(1 + 0(n^{-1/2})).$$

TEMPLE UNIVERSITY, PHILADELPHIA, PENNSYLVANIA.

REFERENCES.

[1] P. Hagis, Jr., "Partitions into odd summands," *American Journal of Mathematics*, vol. 85 (1963), pp. 213-222.

[2] ———, "Partitions into odd and unequal parts," *American Journal of Mathematics*, vol. 86 (1964), pp. 317-324.

[3] S. Iseki, "A partition function with some congruence condition," *American Journal of Mathematics*, vol. 81 (1959), pp. 939-961.

[4] J. Lehner, "A partition function connected with the modulus five," *Duke Mathematical Journal*, vol. 8 (1941), pp. 631-655.

[5] H. Rademacher, "The Fourier coefficient of the modular invariant $J(\tau)$," *American Journal of Mathematics*, vol. 60 (1938), pp. 501-512.

[6] H. Salié, "Zur Abschatzung der Fourierkoeffizienten ganzer Modulformen," *Mathematische Zeitschrift*, vol. 36 (1933), pp. 263-278.

[7] A. Weil, "On some exponential sums," *Proceedings of the National Academy of Sciences, U.S.A.*, vol. 34 (1948), pp. 204-207.

ON THE STRUCTURE OF COMPACT COMPLEX ANALYTIC SURFACES, II.*

By K. KODAIRA.

In Part I of this paper (reference [5]) we have introduced a classification of surfaces. In this Part II we are mainly concerned with surfaces of class VII_0. First, in Section 8, we show that a surface belongs to the class VII if and only if its first Betti number is equal to 1 (Theorem 26). Then, in Section 9, we determine all elliptic surfaces of class VII_0 (Theorem 27). By a Hopf surface we shall mean a surface of which the universal covering manifold is (complex analytically homeomorphic to) $\boldsymbol{C}^2 - O$, where $\boldsymbol{C}^2$ is the space of two complex variables (z_1, z_2) and O denotes the origin $(0, 0)$. In Section 10, we show that all Hopf surfaces belong to the class VII_0 (Theorem 30). We derive some sufficient conditions for a surface to be a Hopf surface (Theorems 34 and 35). We then prove that any deformation of any Hopf surface is a Hopf surface and that any deformation of any elliptic surface of class VII_0 is either an elliptic surface of class VII_0 or a Hopf surface (Theorems 36 and 37). In this connection we remark that, as far as we know, there is no example of surface of class VII_0 other than elliptic and Hopf surfaces. By a volume-preserving complex structure we shall mean a pseudo-group structure corresponding to the pseudo-group of volume-preserving local complex analytic homeomorphisms. In Sections 11 and 12 we examine the application of our results to surfaces with volume-preserving complex structure. We prove that a surface with volume-preserving complex structure is a K3 surface, an Enriques surface or a surface of the form $\boldsymbol{C}^2/G$, where G is a properly discontinuous group of affine transformations of $\boldsymbol{C}^2$ without fixed points (Theorem 39). This result implies that, if a properly discontinuous group $\mathfrak{G}$ of volume-preserving complex analytic automorphisms of $\boldsymbol{C}^2$ is free from fixed points and has a bounded fundamental domain, then $\mathfrak{G}$ is equivalent to a properly discontinuous group of affine transformations of $\boldsymbol{C}^2$ (Theorem 40).

8. Remarks. In the proof of Theorem 21, we used a result that, *if a*

---

[*] This work was partially supported by the National Science Foundation under Grant GP-2606.

Received June 11, 1965.

surface S contains no exceptional curve and if $b_1(S)=1$, $p_g(S)=0$, then S is a minimal surface. First we supply a proof of this result. Suppose that S is not minimal. Then there exists a surface S^* which is birationally equivalent to S and contains exceptional curves C_1 and C_2 with $(C_1C_2)\geqq 1$. Since $(KC_\nu)=(C_\nu{}^2)=-1$, $\nu=1,2,$ the sum $D=C_1+C_2$ is not homologous to zero with respect to rational coefficients and $(D^2)\geqq 0$. This contradicts that $b^+(S^*)=2p_g(S^*)=2p_g(S)=0$.

Next, we prove the following theorem:

THEOREM 26. *A surface belongs to the class* VII *if and only if its first Betti number is equal to* 1.

Proof. Let S be a surface containing no exceptional curve whose Betti number b_1 is equal to 1. If S is non-elliptic, then it is obvious that S belongs to the class VII. Assume that S is elliptic, let Δ be the base curve of S and let Ψ denote the canonical projection: $S\to\Delta$. By Theorem 3, $q=q(S)$ is equal to 1 and $h^{1,0}(S)$ vanishes. It follows that Δ is a projective line. By Theorem 12, the canonical bundle of S is

$$K=\Psi^*(\mathfrak{k}-\mathfrak{f})+\sum(m_\nu-1)[P_\nu].$$

This implies that

$$p_g=\dim|K|+1=\dim|\mathfrak{k}-\mathfrak{f}|+1=\begin{cases}0, & \text{if } c(\mathfrak{f})\geqq 0,\\ -c(\mathfrak{f})-1, & \text{otherwise.}\end{cases}$$

On the other hand, since $q=1$, $c(\mathfrak{f})$ is equal to $-p_g$. Hence p_g vanishes and consequently S belongs to the class VII.

9. Elliptic surfaces of class VII_0.

In this section we denote by Δ a projective line. We fix a system of non-homogeneous coordinates on Δ and indicate a point on Δ by its non-homogeneous coordinate u.

First we let S be an elliptic surface free from singular fibres of which the base curve is Δ. Let Ψ denote the canonical projection of S onto Δ. Obviously all fibres $\Psi^{-1}(u)$, $u\in\Delta$, are complex analytically homeomorphic to one and the same elliptic curve C. We represent C as a quotient group: $C=\boldsymbol{C}/G$, where G is a discontinuous subgroup of the additive group $\boldsymbol{C}$ generated by ω and 1, $\Im\omega>0$, and, for any $\zeta\in\boldsymbol{C}$, we denote by $[\zeta]$ the corresponding element of $C=\boldsymbol{C}/G$. Choose a finite covering $\{E_j\}$ of Δ by small disks E_j. Then $\Psi^{-1}(E_j)$ is a cartesian product: $E_j\times C$ and $(u,[\zeta_j])\in\Psi^{-1}(E_j)$ is identical with $(u,[\zeta_k])\in\Psi^{-1}(E_k)$ if and only if

$$[\zeta_j]=[\epsilon_{jk}\zeta_k+\eta_{jk}(u)], \tag{74}$$

where ϵ_{jk} is a root of unity representing an automorphism: $[\zeta]\to[\epsilon_{jk}\zeta]$ of C and $\eta_{jk}(u)$ is a holomorphic function of u defined on $E_j\cap E_k$. Since Δ is topologically a 2-sphere, the 1-cocycle $\{\epsilon_{jk}\}$ is cohomologous to zero, namely, there exist automorphisms $[\zeta]\to[\epsilon_j\zeta]$ of C such that $\epsilon_{jk}=\epsilon_k/\epsilon_j$. Hence, by the substitution: $\epsilon_j\zeta_j\to\zeta_j$, $\epsilon_j\eta_{jk}(u)\to\eta_{jk}(u)$, the formula (74) is reduced to

$$[\zeta_j]=[\zeta_k+\eta_{jk}(u)]. \tag{75}$$

Thus we see that S is a complex analytic fibre bundle over Δ whose fibre is C and whose structure group is the group of transalations of C. Let $\Omega(C)$ denote the sheaf over Δ of germs of holomorphic functions with values in C. Obviously the structure of S is determined by the cohomology class $\eta\in H^1(\Delta,\Omega(C))$ of the 1-cocycle $\{[\eta_{jk}]\}$, where $[\eta_{jk}]$ denotes the section $u\to[\eta_{jk}(u)]$ of $\Omega(C)$ over $E_j\cap E_k$. Letting Ω denote the sheaf over Δ of germs of holomorphic functions, we have the exact sequence

$$0\to G\to\Omega\to\Omega(C)\to 0.$$

Hence we obtain the isomorphism

$$\delta^*\colon H^1(\Delta,\Omega(C))\to H^2(\Delta,G).$$

We define the characteristic class $c(S)$ of the elliptic surface S to be the image $\delta^*(\eta)\in H^2(\Delta,G)$ of η (see [3], § 11). In view of the isomorphism $H^2(\Delta,G)\cong G$, we may consider $c(S)$ as an element of G. Clearly the elliptic surface S is determined uniquely by its characteristic class $c(S)$. Take a point $a\in\Delta$, $a\neq\infty$, and let E_a denote the disk: $|u-a|<\epsilon$, where ϵ is a positive number. We infer from the above consideration that

$$S=E_a\times C\cup(\Delta-a)\times C, \tag{76}$$

where $(u,[\zeta_a])\in E_a\times C$ *and* $(u,[\zeta])\in(\Delta-a)\times C$ *are identical if and only if*

$$[\zeta]=[\zeta_a+\tfrac{1}{2\pi i}\gamma\log(u-a)],\qquad \gamma\in G.$$

In fact, defining S by the formula (76), we readily verify that the characteristic class $c(S)$ is equal to γ.

Now we examine the structure of elliptic surfaces of class VII_0. Let S be an elliptic surface of class VII_0. S has the following numerical characters: $q=b_1=1$, $h^{1,0}=0$, $p_g=0$, $c_1^2=0$, $b_2=0$, $c_2=0$. The vanishing of $h^{1,0}$ implies that the base curve of S is a projective line Δ. Since the Euler number of an elliptic surface is equal to the sum of Euler numbers of its singular fibres, the vanishing of c_2 implies that any singular fibre of S is of type ${}_mI_0$, $m \geqq 2$. Hence, by a result of Section 4, S is obtained from an elliptic surface S^* free from singular fibres by means of a finite number of logarithmic transformations:

$$S = L_{a_r} L_{a_{r-1}} \cdots L_{a_\nu} \cdots L_{a_2}(S^*),$$

where $a_2, \cdots, a_\nu, \cdots, a_r$ are point on Δ over which S has singular fibres. Obviously we may assume that $a_\nu \neq \infty$. The elliptic surface S^* is written in the form (76). Thus S^* is a logarithmic transform $L_a(\Delta \times C)$ of the cartesian product $\Delta \times C$. Hence, setting $a_1 = a$, we obtain the formula

$$S = L_{a_r} \cdots L_{a_\nu} \cdots L_{a_2} L_{a_1}(\Delta \times C).$$

This formula gives an explicit description of the elliptic surface S. Let m_ν denote the multiplicity of the singular fibre of S over a_ν. Let D_ν be a disk: $|\sigma_\nu| < \epsilon$, where ϵ is a small positive number, and let $\mathfrak{G}_\nu$ denote the cyclic group of order m_ν of complex analytic automorphisms of $D_\nu \times C$ generated by the automorphism

$$g_\nu : (\sigma_\nu, [\zeta_\nu]) \to (\rho_\nu \sigma_\nu, [\zeta_\nu - \beta_\nu]), \qquad \rho_\nu = \exp(2\pi i/m_\nu),$$

where $[\beta_\nu]$ is an element of C of order m_ν. Form the quotient space $D_\nu \times C/\mathfrak{G}_\nu$ and let $((\sigma_\nu, \zeta_\nu))$ denote the point on $D_\nu \times C/\mathfrak{G}_\nu$ corresponding to $(\sigma_\nu, [\zeta_\nu]) \in D_\nu \times C$. Then S is written as a union

$$S = \bigcup_{\nu=1}^{r} D_\nu \times C/\mathfrak{G}_\nu \cup \Delta' \times C, \qquad \Delta' = \Delta - \{a_\nu\}, \tag{77}$$

where $((\sigma_\nu, \zeta_\nu)) \in D_\nu \times C/\mathfrak{G}_\nu$, $\sigma_\nu \neq 0$, is identified with $(u, [\zeta]) \in \Delta' \times C$ if and only if

$$\begin{cases} u = a_\nu + \sigma_\nu{}^{m_\nu}, \\ \zeta \equiv \zeta_\nu + \frac{1}{2\pi i} m_\nu \beta_\nu \log \sigma_\nu \pmod{G}. \end{cases} \tag{78}$$

In order to make explicit that the logarithmic transformation L_{a_ν} defined by (78) involves m_ν and β_ν, we write $L_{a\nu}(m_\nu, \beta_\nu)$ for $L_{a\nu}$.

THEOREM 27. *Any elliptic surface of class* VII$_0$ *can be written in the form*

$$S = L_{a_r}(m_r, \beta_r) \cdots L_{a_\nu}(m_\nu, \beta_\nu) \cdots L_{a_1}(m_1, \beta_1)(\Delta \times C), \tag{79}$$

where $a_\nu \neq a_\lambda$ *for* $\nu \neq \lambda$ *and*

$$\beta_1 + \cdots + \beta_\nu + \cdots + \beta_r \neq 0. \tag{80}$$

Proof. It suffices to show that the elliptic surface S defined by (79) belongs to the class VII_0 if and only if the inequality (80) holds. Since the Chern numbers c_1^2 and c_2 of S vanish, we get, using (3),

$$p_g - q + 1 = 0.$$

Hence we infer from Theorem 12 that

$$K = -2C_u + \sum_{\nu=1}^{r} (m_\nu - 1) P_\nu \qquad P_\nu = \Psi^{-1}(a_\nu),$$

is a canonical divisor on S, where C_u is a general fibre and Ψ denotes the canonical projection of S onto Δ. It follows that $p_g = 0$ and $q = 1$. Hence, by Theorem 3, the first Betti number b_1 of S is equal to 1 or 2 according as $h^{1,0}$ is equal to 0 or 1. For our purpose it suffices therefore to show that there exists on S a holomorphic 1-form $\phi \neq 0$ if and only if $\sum \beta_\nu = 0$.

Assume the existence of a holomorphic 1-form $\phi \neq 0$ on S. Letting $(u, [\zeta])$ denote a "variable" point on $\Delta' \times C$, we write

$$\phi = f(u)\,du + h\,d\zeta, \qquad \text{on } \Delta' \times C.$$

The coefficients $f(u)$ and h are holomorphic functions of $u \in \Delta'$, since they are constant on each fibre $u \times C$, while, by Theorem 1, $d\phi$ vanishes. Hence h is reduced to a constant. By the substitution (78) we get

$$\phi = m_\nu (f(u)\sigma_\nu^{m_\nu} + \tfrac{1}{2\pi i} h\beta_\nu)\sigma_\nu^{-1} d\sigma_\nu + h\,d\zeta_\nu.$$

Hence $f(u)\sigma_\nu^{m_\nu} + \frac{1}{2\pi i} h\beta_\nu$ is (extended to) a holomorphic function of σ_ν, $|\sigma_\nu| < \epsilon$, which vanishes at $\sigma_\nu = 0$. It follows that $f(u)$ has the Laurent expansion

$$f(u) = -\frac{1}{2\pi i} \cdot \frac{h\beta_\nu}{(u - a_\nu)} + \cdots, \qquad \text{at } a_\nu.$$

On the other hand, $u^2 f(u)$ is holomorphic at $u = \infty$, since the differential form $f(u)du$ is holomorphic at $u = \infty$. Hence

$$f(u) = -\frac{1}{2\pi i} h \sum_{\nu=1}^{r} \frac{\beta_\nu}{(u - a_\nu)}, \qquad h \neq 0,$$

and

$$\sum_{\nu=1}^{r} \beta_\nu = -2\pi i h^{-1} \lim_{u \to \infty} u f(u) = 0.$$

Conversely, if $\sum \beta_\nu = 0$, then

$$\phi = -\frac{1}{2\pi i}\sum_{\nu=1}^{r}\frac{\beta_\nu}{(u-a_\nu)}\,du + d\zeta$$

is a holomorphic 1-form on S, q. e. d.

Since the logarithmic transformations in (79) are computative, we may assume that $m_1 = \cdots = m_{\lambda-1} = 1$ and $m_\nu \geqq 2$ for $\nu \geqq \lambda$. Let

$$S^* = L_{a_{\lambda-1}}(1, \beta_{\lambda-1}) \cdots L_{a_1}(1, \beta_1)(\Delta \times C).$$

The elliptic surface S^* is free from singular fibres. Hence it can be written in the form (76), namely,

$$S^* = L_a(1, \gamma)(\Delta \times C), \qquad \gamma \in G.$$

Consider the case in which $r \geqq \lambda$. Since a is an arbitrary point on Δ, we may assume that $a = a_\lambda$. We then have

$$L_{a_\lambda}(m_\lambda, \beta_\lambda)(S^*) = L_{a_\lambda}(m_\lambda, \beta_\lambda + \gamma)(\Delta \times C).$$

Consequently, writing β_λ in place of $\beta_\lambda + \gamma$, we obtain

$$S = L_{a_r}(m_r, \beta_r) \cdots L_{a_\lambda}(m_\lambda, \beta_\lambda)(\Delta \times C).$$

Thus we conclude that *the representation* (79) *of S can be normalized in the following manner*: i) *In the case in which S has singular fibres, r is equal to the number of singular fibres and $m_\nu \geqq 2$ for $\nu = 1, 2, \cdots, r$.* ii) *In the case in which S is free from singular fibres, r is equal to* 1 *and $m_1 = 1$.*

Next, we examine the universal covering manifolds of elliptic surfaces of class VII_0. We denote by $\boldsymbol{W}$ the domain $\boldsymbol{C}^2 - O$, where $\boldsymbol{C}^2$ is the space of two complex variables (z_1, z_2) and O denotes the origin $(0, 0)$.

LEMMA 8. *Let S be an elliptic surface of class* VII_0. *If S has not more than two singular fibres, then the universal covering manifold of S is* (*complex analytically homeomorphic to*) *the domain $\boldsymbol{W} = \boldsymbol{C}^2 - O$.*

Proof. In view of the above results, S can be written in the form

$$\begin{aligned} S &= L_{a_2}(m_2, \beta_2) L_{a_1}(m_1, \beta_1)(\Delta \times C) \\ &= D_1 \times C/\mathfrak{G}_1 \cup D_2 \times C/\mathfrak{G}_2 \cup \Delta' \times C, \qquad \Delta' = \Delta - \{a_1, a_2\}. \end{aligned}$$

Hence, assuming that $a_1 = 0$, $a_2 = \infty$, and letting D_1 and D_2 be copies of $\boldsymbol{C}$, we obtain

$$S = D_1 \times C/\mathfrak{G}_1 \cup D_2 \times C/\mathfrak{G}_2, \tag{81}$$

where $((\sigma_1, \zeta_1)) \in D_1 \times C/\mathfrak{G}_1$ and $((\sigma_2, \zeta_2)) \in D_2 \times C/\mathfrak{G}_2$ are identical if and only if

$$(82) \qquad \begin{cases} \sigma_1{}^{m_1}\sigma_2{}^{m_2} = 1, \\ \zeta_1 + \frac{1}{2\pi i} m_1\beta_1 \log \sigma_1 \equiv \zeta_2 + \frac{1}{2\pi i} m_2\beta_2 \log \sigma_2 \,(\mathrm{mod}\, G). \end{cases}$$

We note that the canonical projection of S onto Δ is

$$\Psi : ((\sigma_1, \zeta_1)) \to u = \sigma_1{}^{m_1}, \quad ((\sigma_2, \zeta_2)) \to u = \sigma_2{}^{-m_2}.$$

First we consider the case in which m_1 and m_2 are relatively prime. We fix integers p and q such that

$$pm_1 + qm_2 = 1.$$

$m_1m_2(\beta_1 + \beta_2)$ belongs to G. Hence, by an appropriate choice of the coordinate ζ on $\boldsymbol{C}$, we may assume that

$$m_1m_2(\beta_1 + \beta_2) = n$$

is a positive integer. We have

$$m_1(m_2\beta_2 - np) + m_2(m_1\beta_1 - nq) = 0.$$

Hence we find an element γ of G such that

$$m_1\beta_1 - nq = m_1\gamma, \qquad m_2\beta_2 - np = - m_2\gamma.$$

It follows that $[\beta_1] = [nq/m_1]$, $[\beta_2] = [np/m_2]$. Moreover the formula (82) is reduced to

$$(82)' \qquad \begin{cases} \sigma_1{}^{m_1}\sigma_2{}^{m_2} = 1, \\ \zeta_1 + \frac{1}{2\pi i} nq \log \sigma_1 \equiv \zeta_2 + \frac{1}{2\pi i} np \log \sigma_2 \,(\mathrm{mod}\, G). \end{cases}$$

Consequently we may replace β_1 and β_2 by nq/m_1 and np/m_2, respectively.

Setting $\alpha_\nu = \exp(2\pi i\, m_\nu\omega/n)$, $\epsilon_\nu = \exp(2\pi i\, m_\nu/n)$, for $\nu = 1, 2$, we define linear transformations

$$A : (z_1, z_2) \to (\alpha_1 z_1, \alpha_2 z_2), \qquad B : (z_1, z_2) \to (\epsilon_1 z_1, \epsilon_2 z_2).$$

Since n and m_ν are relatively prime, ϵ_ν is a primitive n-th root of unity. Moreover $|\alpha_\nu|$ is smaller than 1. Hence A and B generate a discontinuous group $\mathfrak{G}$ of complex analytic automorphisms without fixed points of the domain $\boldsymbol{W}$ and the quotient space $\boldsymbol{W}/\mathfrak{G}$ is a (compact complex analytic) surface. We infer readily that the following formula defines a biholomorphic map μ of $\boldsymbol{W}/\mathfrak{G}$ onto S:

$$\mu(z_1, z_2) = \begin{cases} ((\sigma_1, \zeta_1)), & \text{if } z_1 \neq 0, \\ ((\sigma_2, \zeta_2)), & \text{if } z_2 \neq 0, \end{cases}$$

where

$$\sigma_1 = z_2 \exp(-(m_2/m_1)\log z_1), \qquad \zeta_1 = \tfrac{1}{2\pi i}(n/m_1)\log z_1,$$
$$\sigma_2 = z_1 \exp(-(m_1/m_2)\log z_2), \qquad \zeta_2 = \tfrac{1}{2\pi i}(n/m_2)\log z_2.$$

Thus S is complex analytically homeomorphic to $\boldsymbol{W}/\mathfrak{G}$ and consequently $\boldsymbol{W}$ is the universal covering manifold of S.

In the case in which m_1 and m_2 are not relatively prime, we let d denote the greatest common divisor of m_1 and m_2. We take a projective line Γ with a system of non-homogeneous coordinate v and consider the analytic fibre space S_Γ over Γ which is induced from S by the map: $v \to u = v^d$ of Γ onto Δ (see [3], § 6). Letting $\{g_\nu{}^d\}$ denote the subgroup of $\mathfrak{g}_\nu$ generated by $g_\nu{}^d$, we infer readily that

$$\begin{aligned} S_\Gamma &= D_1 \times C/\{g_1{}^d\} \cup D_2 \times C/\{g_2{}^d\} \\ &= L_0(m_1/d, \beta_1 d) L_\infty(m_2/d, \beta_2 d)(\Gamma \times C). \end{aligned}$$

Moreover S_Γ is an unramified covering surface of S. By the above result, the universal covering manifold of S_Γ is $\boldsymbol{W}$. Consequently the universal covering manifold of S is $\boldsymbol{W}$, q.e.d.

Consider an elliptic surface S of class VII$_0$ with a normalized representation:

$$S = L_{a_r}(m_r, \beta_r) \cdots L_{a_\nu}(m_\nu, \beta_\nu) \cdots L_{a_1}(m_1, \beta_1)(\Delta \times C).$$

It follows from Theorem 12 that

$$K = -2C_u + \sum_{\nu=1}^{r} (m_\nu - 1) P_\nu, \qquad P_\nu = \Psi^{-1}(a_\nu),$$

is a canonical divisor on S, where C_u denotes a general fibre of S. We define

$$\kappa = \kappa(S) = -2 + \sum_{\nu=1}^{r} (1 - 1/m_\nu).$$

We denote by $\boldsymbol{D}$ the unit disk: $|z| < 1$ in the space $\boldsymbol{C}$ of a complex variable z.

THEOREM 28. *The universal covering manifold $\mathfrak{U}$ of S is given by the following formula:*

$$\mathfrak{U} = \begin{cases} \boldsymbol{D} \times \boldsymbol{C}, & \text{if } \kappa(S) > 0, \\ \boldsymbol{C}^2, & \text{if } \kappa(S) = 0, \\ \boldsymbol{W} & \text{if } \kappa(S) < 0. \end{cases} \tag{83}$$

Proof. i) If S has not more than two singular fibres, then $\kappa(S)$ is obviously negative and, by Lemma 8, $\mathfrak{U}$ is (complex analytically homeomorphic to) $\boldsymbol{W}$.

ii) Assume that S has at least three singular fibres, i.e., $r \geqq 3$ and $m_\nu \geqq 2$ for $\nu = 1, 2, \cdots, r$. There exists a simply connected covering Riemann surface Γ of Δ which is unramified over $\Delta - \{a_\nu\}$ and has branch points of order $m_\nu - 1$ over each point a_ν. Let ϖ denote the covering map of Γ onto Δ. We note that, in a neighborhood of a point $b \in \Gamma$, the covering map ϖ can be written in terms of an appropriate local uniformization variable v_b of the center b on Γ in the form

$$\varpi : v_b \to u = \begin{cases} a_\nu + v_b^{m_\nu} & \text{if } \varpi(b) = a_\nu, \\ 1/v_b & \text{if } \varpi(b) = \infty, \\ \varpi(b) + v_b & \text{otherwise.} \end{cases}$$

Let S_Γ be the analytic fibre space of elliptic curves over Γ which is induced from S by the map $\varpi : \Gamma \to \Delta$. Obviously S_Γ is free from singular fibres and forms an unramified covering manifold of S. Hence $\mathfrak{U}$ coincides with the universal covering manifold of S_Γ.

If $\kappa = \kappa(S) > 0$, then Γ is conformally equivalent to $\boldsymbol{D}$. Hence the fibre space S_Γ is trivial, i.e., S_Γ is (complex analytically homeomorphic to) $\boldsymbol{D} \times C$. Obviously the universal covering manifold $\mathfrak{U}$ of $S_\Gamma = \boldsymbol{D} \times C$ is $\boldsymbol{D} \times \boldsymbol{C}$.

If $\kappa = 0$, then Γ is conformally equivalent to $\boldsymbol{C}$ and therefore S_Γ is $\boldsymbol{C} \times C$. Consequently $\mathfrak{U}$ is $\boldsymbol{C}^2$.

If $\kappa < 0$, then Γ is conformally equivalent to a projective line and therefore S is an elliptic surface free from singular fibres. Hence S_Γ can be written in the form

$$S_\Gamma = L_0(1, \gamma)(\Gamma \times C), \qquad \gamma \in G.$$

Since S is non-algebraic, S_Γ is non-algebraic and therefore $\gamma \neq 0$. Hence S_Γ is an elliptic surface of class VII_0 which has no singular fibre. Consequently, by Lemma 8, the universal covering manifold $\mathfrak{U}$ of S_Γ is $\boldsymbol{W}$, q. e. d.

The elliptic surface

$$S = L_{a_r}(m_r, \beta_r) \cdots L_{a_\nu}(m_\nu, \beta_\nu) \cdots L_{a_1}(m_1, \beta_1)(\Delta \times C) \tag{84}$$

depends holomorphically on the parameters $a_1, \cdots, a_r, \omega$. We consider the case in which $r \geqq 3$ and $m_\nu \geqq 2$ for $\nu = 1, \cdots, r$. The constants β_ν in (84) are linear forms in ω:

$$m_\nu \beta_\nu = h_\nu \omega + k_\nu,$$

where the coefficients h_ν and k_ν are rational integers. We fix the integers $m_\nu, h_\nu, k_\nu, \nu = 1, 2, \cdots, r$. The complex structure of S is invariant under any linear fractional transformation of Δ. Hence we fix the points a_1, a_2, a_3 on Δ and write $S_{a_4, \cdots, a_r, \omega}$ for S. Let $\boldsymbol{H}$ denote the upper half ω-plane and let B be the open subset of $\Delta \times \cdots \times \Delta \times \boldsymbol{H}$ consisting of those points $(a_4, \cdots, a_\nu, \cdots, a_r, \omega)$ which satisfy the condition that $a_\nu \neq a_1, \neq a_2, \neq a_3$, $a_\lambda \neq a_\nu$ for $\lambda \neq \nu$. It is clear that the set

$$\mathcal{F} = \{S_{a_4, \cdots, a_r, \omega} \mid (a_4, \cdots, a_r, \omega) \in B\}$$

forms a complex analytic family.

THEOREM 29. *The complex analytic family* $\mathcal{F}$ *consisting of elliptic surfaces* $S_{a_4, \cdots, a_r, \omega}$, $(a_4, \cdots, a_r, \omega) \in B$, *is effectively parametrized and complete.*

Proof. It suffices to consider the case in which $a_\nu \neq \infty$ for $\nu = 1, 2, \cdots, r$. We denote the infinitesimal deformations of $S = S_{a_4, \cdots, a_r, \omega}$ along the tangent vectors $\partial/\partial a_\nu$ and $\partial/\partial\omega$ by the symbols $\partial S/\partial a_\nu$ and $\partial S/\partial\omega$, respectively. In view of a theorem of completeness (see Kodaira and Spencer [8]), it suffices for our purpose to prove that the infinitesimal deformations $\partial S/\partial a_4, \cdots, \partial S/\partial a_r$, $\partial S/\partial\omega$ form a base of $H^1(S, \Theta)$, where Θ denotes the sheaf over S of germs of holomorphic vector fields. We write

$$S = \bigcup_{\nu=1}^{r} U_\nu \cup \Delta' \times C, \qquad U_\nu = D_\nu \times C/\mathfrak{G}_\nu, \qquad \Delta' = \Delta - \{a_\nu\}$$

(see (77)). We employ (σ_ν, ζ_ν) and (u, ζ) as local coordinates of points $((\sigma_\nu, \zeta_\nu))$ on U_ν and $(u, [\zeta])$ on $\Delta' \times C$, respectively. We derive from (78) the formulae

$$\begin{cases} \partial/\partial\sigma_\nu = m_\nu \sigma_\nu^{m_\nu - 1} \partial/\partial u - \frac{1}{2\pi i} m_\nu \beta_\nu \sigma_\nu^{-1} \partial/\partial\zeta, \\ \partial/\partial\zeta_\nu = \partial/\partial\zeta. \end{cases} \tag{85}$$

First we calculate the dimension of $H^0(S, \Theta)$. Take a holomorphic vector field $\theta \in H^0(S, \Theta)$ and write

$$\theta = \theta_1(u)\partial/\partial u + \theta_2(u)\partial/\partial\zeta, \qquad \text{on } \Delta' \times C.$$

The coefficient $\theta_1(u)$ is a holomorphic function of $u \in \Delta'$, $u \neq \infty$, and, since $\theta_1(u)\partial/\partial u$ is holomorphic at $u = \infty$, $\theta_1(u)$ has a pole of order $\leqq 2$ at $u = \infty$. The coefficient $\theta_2(u)$ is holomorphic on Δ'. Using (85) we get

$$\theta = m_\nu^{-1} \sigma_\nu^{1 - m_\nu} \theta_1(u)\, \partial/\partial\sigma_\nu + (\theta_2(u) + \tfrac{1}{2\pi i} \beta_\nu \sigma_\nu^{-m_\nu} \theta_1(u))\partial/\partial\zeta_\nu. \tag{86}$$

Hence $\sigma_\nu^{1-m_\nu}\theta_1(a_\nu+\sigma_\nu^{m_\nu})$ is (extended to) a holomorphic function of σ_ν, $|\sigma_\nu|<\epsilon$, and therefore $\theta_1(a_\nu)$ vanishes for $\nu=1,2,3,\cdots$. Consequently $\theta_1(u)$ vanishes identically. Now (86) shows that $\theta_2(u)$ is holomorphic at a_ν. Therefore $\theta_2(u)$ is reduced to a constant. Thus we see that

$$\dim H^0(S,\Theta)=1. \tag{87}$$

By the duality theorem we have

$$\dim H^2(S,\Theta)=\dim H^0(S,\Omega^1\otimes\Omega^2),$$

where $\otimes$ denotes the tensor product over $\mathcal{O}$. Take an element ϕ of $H^0(S,\Omega^1\otimes\Omega^2)$ and write

$$\phi=(f(u)\,du+h(u)\,d\zeta)\otimes du\wedge d\zeta, \qquad \text{on } \Delta'\times C.$$

By the substitution (78) we get

$$\begin{aligned}\phi=(m_\nu^2\sigma_\nu^{2m_\nu-2}f(u)\,d\sigma_\nu&+\tfrac{1}{2\pi i}m_\nu^2\beta_\nu\sigma_\nu^{m_\nu-2}h(u)\,d\sigma_\nu\\&+m_\nu\sigma_\nu^{m_\nu-1}h(u)\,d\zeta_\nu)\otimes d\sigma_\nu\wedge d\zeta_\nu.\end{aligned}$$

Hence we infer that $h(u)$ and $(u-a_\nu)f(u)$ are holomorphic at $u=a_\nu$. Since $h(u)\,du$ is holomorphic at $u=\infty$, $h(\infty)$ vanishes. Consequently $h(u)$ vanishes identically. Since $f(u)\,du\otimes du$ is holomorphic at $u=\infty$, $f(u)$ has a zero of order $\geqq 4$ at $u=\infty$. Therefore $f(u)$ is a rational function of the form

$$f(u)=\sum_{\nu=1}^{r}k_\nu(u-a_\nu)^{-1}$$

and the coefficients k_ν satisfy the simultaneous linear equations

$$\sum_{\nu=1}^{r}k_\nu=\sum_{\nu=1}^{r}a_\nu k_\nu=\sum_{\nu=1}^{r}a_\nu^2k_\nu=0.$$

Thus we see that

$$\dim H^2(S,\Theta)=r-3. \tag{88}$$

The Chern numbers c_1^2 and c_2 of S vanish. Hence, by the Riemann-Roch-Hirzebruch theorem we have

$$\dim H^1(S,\Theta)=\dim H^0(S,\Theta)+\dim H^2(S,\Theta).$$

Combining this with (87) and (88), we obtain

$$\dim H^1(S,\Theta)=r-2. \tag{89}$$

The projection $\Psi(U_\nu)$ of U_ν is a disk on Δ of radius $\epsilon^{m\nu}$ and of center a_ν. Let Δ_0 denote the open subset of Δ defined by the inequalities: $|u - a_\nu| > \frac{1}{2}\epsilon^{m\nu}$, $\nu = 1, 2, \cdots, r$, and let $U_0 = \Delta_0 \times C$. Obviously the set $\mathfrak{U} = \{U_0, U_1, \cdots, U_r\}$ forms a covering of S. Since $U_\nu \cap U_\lambda$ is empty for $1 \leqq \nu < \lambda \leqq r$, any 1-cocycle on the nerve of $\mathfrak{U}$ with coefficients in Θ is composed of holomorphic vector fields $\theta_{0\lambda}$ and $\theta_{\lambda 0} = -\theta_{0\lambda}$, $\lambda = 1, 2, \cdots, r$, defined respectively on $U_0 \cap U_\lambda$. We shall denote the 1-cocycle by $\{\theta_{0\lambda}\}$. We infer from (78) that the infinitesimal deformation $\partial S/\partial a_\nu$ is the cohomology class of the 1-cocycle

$$\theta^{(\nu)} = \{\theta^{(\nu)}{}_{0\lambda}\}, \qquad \theta^{(\nu)}{}_{0\lambda} = \delta_{\lambda\nu}\partial/\partial u. \tag{90}$$

To evaluate $\partial S/\partial\omega$, we cover C by disks V_j, $j = 1, 2, \cdots, n$, and form a refinement $\mathfrak{W} = \{W_1, \cdots, W_j, \cdots\}$ of $\mathfrak{U}$ such that

$$\begin{aligned} &W_j = \Delta_0 \times V_j, && \text{for } j = 1, 2, \cdots, n \\ &W_j \subset U_\nu,\ \nu \geqq 1, && \text{for } j = n+1, n+2, \cdots, \end{aligned}$$

where the index ν depends on j. We fix a branch ζ_j of ζ on each disk V_j and employ (u, ζ_j) as local coordinates on $\Delta_0 \times V_j$. We have

$$\zeta_j = \zeta_k + m_{jk}\omega + n_{jk}, \qquad m_{jk}, n_{jk} \in \mathbf{Z}.$$

Hence we infer that $\partial S/\partial\omega$ is the cohomology class of a 1-cocycle $\eta = \{\eta_{jk}\}$ on the nerve of $\mathfrak{W}$ such that

$$\eta_{jk} = m_{jk}\partial/\partial\zeta, \qquad \text{for } j, k = 1, 2, \cdots, n. \tag{91}$$

Let $\{\theta^{(\nu)}{}_{jk}\}$ denote the 1-cocycle on the nerve of $\mathfrak{W}$ which is the image of $\theta^{(\nu)}$ under the simplicial map $\mathfrak{W} \to \mathfrak{U}$. It is clear that

$$\theta^{(\nu)}{}_{jk} = 0, \qquad \text{for } j, k = 1, 2, \cdots, n. \tag{92}$$

In view of (89) it suffices for our purpose to show that $\partial S/\partial a_4, \cdots, \partial S/\partial a_r$, $\partial S/\partial\omega$ are linearly independent. Assume a cohomological relation

$$\sum_{\nu=4}^{r} c_\nu\theta^{(\nu)} + c\eta \sim 0. \tag{93}$$

This implies the existence of a 0-cochain $\{\theta_j\}$ composed of holomorphic vector fields θ_j on W_j such that

$$\sum_{\nu=4}^{r} c_\nu\theta^{(\nu)}{}_{jk} + c\eta_{jk} = \theta_k - \theta_j.$$

Combining this with (91), (92) and letting

$$\theta_j = A_j(u, \zeta_j)\partial/\partial u + B_j(u, \zeta_j)\partial/\partial \zeta_j, \qquad \text{for } j = 1, \cdots, n,$$

we obtain

$$c m_{jk}\partial/\partial\zeta = B_k(u, \zeta_k)\partial/\partial\zeta_k - B_j(u, \zeta_j)\partial/\partial\zeta_j.$$

On the other hand the 1-cocycle $\{m_{jk}\partial/\partial\zeta\}$ represents the infinitesimal deformation $\partial C_u/\partial\omega$ of the fibre C_u. Hence the constant c vanishes. (93) implies therefore the existence of a 0-cochain $\{\theta_\lambda\}$ composed of holomorphic vector fields θ_λ on U_λ such that

$$\sum_{\nu=4}^{r} c_\nu \theta^{(\nu)}{}_{0\lambda} = \theta_\lambda - \theta_0.$$

Combining this with (90) and setting

$$\theta_0 = A_0(u)\partial/\partial u + B_0(u)\partial/\partial\zeta.$$
$$\theta_\nu = A_\nu(\sigma_\nu)\partial/\partial\sigma_\nu + B_\nu(\sigma_\nu)\partial/\partial\zeta_\nu \qquad \nu = 1, \cdots, r,$$

we infer with the aid of (85) that

$$m_\nu \sigma_\nu^{m_\nu - 1} A_\nu(\sigma_\nu) - A_0(u) = \begin{cases} 0, & \text{for } \nu = 1, 2, 3, \\ c_\nu, & \text{for } \nu = 4, \cdots, r. \end{cases}$$

This proves that $A_0(u)$ is holomorphic for $u \neq \infty$ and vanishes at $u = a_1, a_2, a_3$, while $A_0(u)$ has a pole of order $\leqq 2$ at $u = \infty$. Hence $A_0(u)$ vanishes identically and consequently $c_\nu = 0$. Thus we conclude that the cohomological relation (93) is trivial, q. e. d.

10. Hopf surfaces. *Definition. By a Hopf surface we shall mean a surface of which the universal covering manifold is (complex analytically homeomorphic to) the domain* $\boldsymbol{W} = \boldsymbol{C}^2 - O$. *By a primary Hopf surface we shall mean a Hopf surface whose fundamental group is infinite cyclic.*

Consider a Hopf surface $S = \boldsymbol{W}/G$, where G denotes the covering transformation group of $\boldsymbol{W}$ with respect to S. It follows from a theorem of Hartogs that any complex analytic automorphism of $\boldsymbol{W}$ can be extended to a complex analytic automorphism of $\boldsymbol{C}^2$. Hence we may assume that each element $g \in G$ is a complex analytic automorphism of $\boldsymbol{C}^2$ which leaves invariant the origin O.

Let B denote the closed spherical set in $\boldsymbol{C}^2$ determined by the inequality: $|z_1|^2 + |z_2|^2 \leqq 1$. A complex analytic automorphism g of $\boldsymbol{C}^2$ is called a *contraction* if $g^n(B)$ converges to O for $n \to +\infty$. *The group G contains a contraction.* We verify this as follows: Let ∂B denote the (set theoretical) boundary of B. Choose an element $g \in G$ such that $g(\partial B)$ does not meet ∂B. Then, since g leaves invariant the origin O, either $g(B)$ is contained in

$B-\partial B$ or B is contained in $g(B-\partial B)$. We may assume the first alternative, since the replacement of g by g^{-1} reduces the second alternative to the first one. We have

$$g^n(B)\subset g^{n-1}(B)-\partial g^{n-1}(B), \qquad \text{for } n=1,2,\cdots.$$

Our purpose is to show that $\bigcap_n g^n(B)=O$. Suppose that $z\neq O$ is a point on the boundary of $\bigcap_n g^n(B)$ and let N be a small neighborhood of z. It is clear that $N\not\subset g^n(B)$ for all sufficiently large values of n, while z is an inner point of $g^n(B)$. Hence N meets $g^n(\partial B)$ for all sufficiently large values of n. This contradicts the proper discontinuity of G. Thus we see that g is a contraction.

For any contraction g, we let $\{g\}$ denote the infinite cyclic group generated by g. Obviously the quotient space $W/\{g\}$ is a (compact complex analytic) surface. *If $g\in G$ is a contraction, then there exists a positive integer n such that g^n belongs to the center of G.* We verify this as follows: Since $W/\{g\}$ forms a finite covering surface of $S=W/G$, the index of $\{g\}$ in G is finite. Hence $\{g\}$ contains a normal subgroup of G of finite index which is obviously an infinite cyclic group $\{g^n\}$ generated by a certain power g^n of g. For an arbitrary element $h\in G$, hg^nh^{-1} is a generator of $\{g^n\}=h\{g^n\}h^{-1}$, while g^n and hg^nh^{-1} are both contractions. Hence hg^nh^{-1} coincides with g^n. Thus g^n belongs to the center of G.

Take a contraction f belonging to the center of G and let

$$\tilde{S}=W/\{f\}.$$

The primary Hopf surface $\tilde{S}$ thus defined is a finite unramified normal covering surface of S. In case the group G is infinite cyclic, G is generated by a contraction. Hence *any primary Hopf surface has the form* $W/\{f\}$.

In view of a classical result (see Lattès [9] and Sternberg [10]), we may assume by an appropriate choice of the system of (global) complex coordinates (z_1,z_2) on C^2 that the contraction f has the *normal form*

$$f:(z_1,z_2)\rightarrow(z_1',z_2')=(\alpha_1z_1+\lambda z_2^m,\alpha_2z_2), \tag{94}$$

where m is a positive integer and the constants α_1, α_2, λ are subject to the restrictions

$$(\alpha_1-\alpha_2^m)\lambda=0, \qquad 0<|\alpha_1|\leqq|\alpha_2|<1. \tag{95}$$

Note that the characteristic values α_1 and α_2 are determined uniquely by f.

Let S^n denote a differentiable n-sphere. We infer from (94) that $\tilde{S} = W/\{f\}$ is differentiably homeomorphic to $S^1 \times S^3$. Since $\tilde{S}$ is a finite unramified covering surface of S, it follows that $b_1(S) = 1$ and $b_2(S) = 0$. The vanishing of $b_2(S)$ implies that S contains no exceptional curve. Hence, by Theorem 26, S belongs to the class VII_0. Thus we obtain the following

THEOREM 30. *Any Hopf surface belongs to the class* VII_0 *and has a finite unramified normal covering surface which is a primary Hopf surface. Any primary Hopf surface is a quotient space* $W/\{f\}$ *of* W *with respect to an infinite cyclic group* $\{f\}$ *generated by a complex analytic automorphism* f *of the form* (94).

Now we consider a primary Hopf surface $S = W/\{f\}$, where f is a contraction of the form (94). The equation: $z_2 = 0$ is invariant under f and defines a non-singular elliptic curve C on S which is isomorphic to $C^*/\{\alpha_1\}$, where $\{\alpha_1\}$ denotes the infinite cyclic subgroup of C^* generated by α_1. In the case where $\lambda = 0$, the equation: $z_1 = 0$ is also invariant under f and defines another non-singular elliptic curve C_1 of S. There exists on S a meromorphic 2-form ψ defined by the formula

$$(96) \qquad \psi = \begin{cases} z_2^{-m-1}\, dz_1 \wedge dz_2, & \text{for } \lambda \neq 0, \\ z_1^{-1} z_2^{-1}\, dz_1 \wedge dz_2, & \text{for } \lambda = 0. \end{cases}$$

Hence the canonical bundle of S is

$$(97) \qquad K = \begin{cases} -(m+1)[C], & \text{for } \lambda \neq 0, \\ -[C_1 + C], & \text{for } \lambda = 0. \end{cases}$$

The normal bundle of C in S is the restriction $[C]_C$ of $[C]$ to C. It follows from (94) that

$$(98) \qquad [C]_C = C^* \times C/\{f^*\},$$

where $\{f^*\}$ denotes the infinite cyclic group generated by the automorphism $f^*: (z_1, \zeta) \to (\alpha_1 z_1, \alpha_2 \zeta)$ of $C^* \times C$. The Picard variety attached to C is isomorphic to C. We infer from (98) that $[C]_C$ *is the element of the Picard variety* $C = C^*/\{\alpha_1\}$ *which corresponds to* $\alpha_2 \in C^*$ *by means of the natural homomorphism*: $C^* \to C^*/\{\alpha_1\}$. In the case in which $\alpha_2 \neq \alpha_1$, it follows that $[C]_C$ is non-trivial and that

$$(99) \qquad \dim H^1(C, \mathcal{O}([C]_C)) = 0.$$

THEOREM 31. $W/\{f\}$ *is an elliptic surface if and only if* $\lambda = 0$ *and* $\alpha_1^k = \alpha_2^l$ *for certain positive integers* k, l.

Proof. Let $S=W/\{f\}$. If $\lambda=0$ and if $\alpha_1{}^k=\alpha_2{}^l$, then $z_1{}^k/z_2{}^l$ is invariant under f and represents a non-constant meromorphic function on S. Hence S is elliptic.

If, conversely, S is elliptic, then the elliptic curve C is a fibre of S. Hence there exists on S a non-constant meromorphic function x whose polar cycle is a multiple nC of C. $\phi(z_1,z_2)=z_2{}^n x$ is (extended to) a holomorphic function on C^2 and satisfies the functional equation

$$(100)\qquad \phi(\alpha_1 z_1+\lambda z_2{}^m,\alpha_2 z_2)=\alpha_2{}^n\phi(z_1,z_2).$$

Letting

$$\phi_\nu=\phi_\nu(z_1,z_2)=\partial^\nu\phi(z_1,z_2)/\partial z_1{}^\nu,$$

we have

$$(101)_\nu\qquad \alpha_1{}^\nu\phi_\nu(\alpha_1 z_1+\lambda z_2{}^m,\alpha_2 z_2)=\alpha_2{}^n\phi_\nu(z_1,z_2).$$

Hence we infer that ϕ_ν vanishes identically for sufficiently large values of ν. Let k be the minimum integer such that ϕ_{k+1} vanishes identically. If $k=0$, then it would follow from (100) that $\phi(z_1,z_2)=cz_2{}^n$, where c is a constant. This contradicts that x is non-constant. Thus k is a positive integer. $\phi_k=\phi_k(z_2)$ is independent of z_1 and satisfies

$$\alpha_1{}^k\phi_k(\alpha_2 z_2)=\alpha_2{}^n\phi_k(z_2).$$

Hence we get

$$\alpha_1{}^k=\alpha_2{}^l,\qquad \phi_k(z_2)=az_2{}^{n-l},\qquad a\neq 0,$$

where l is a positive integer, and

$$\phi_{k-1}(z_1,z_2)=az_1z_2{}^{n-l}+\sum a_h z_2{}^h.$$

Inserting this in $(101)_{k-1}$, we obtain

$$a\lambda=a_{m+n-l}(\alpha_1-\alpha_2{}^m).$$

Combining this with (95), we conclude that $\lambda=0$, q. e. d.

We shall examine the covering transformation group G of a non-elliptic Hopf surface W/G. Take a contraction f in the center of G and represent it in the form (94). Then the primary Hopf surface $W/\{f\}$ is non-elliptic and therefore, by Theorem 31, either $\lambda\neq 0$ or $\alpha_1{}^k\neq\alpha_2{}^h$ for any positive integers k and h. Let g be an arbitrary element of G.

In the case in which $\lambda\neq 0$, we infer from the relation $gf=fg$ that g is a transformation of the form

$$g:(z_1,z_2)\to(\beta^m z_1+\mu z_2{}^m,\beta z_2),\qquad \beta\neq 0.$$

Hence G is abelian. If $\beta^m = 1$, then $(z_1, 0)$ is a fixed point of g and therefore $g = 1$. Hence $g \to \beta^m$ is an isomorphism of G onto a discontinuous subgroup of the multiplicative group $\boldsymbol{C}^*$. It follows that G is a direct product $Z \times Z_l$ of an infinite cyclic group Z and a finite cyclic group Z_l of order l. It is clear that, if $g \in Z_l$, then the coefficient μ vanishes.

In the case in which $\lambda = 0$, we infer from the relation $gf = fg$ that g is a linear transformation of the form

$$g : (z_1, z_2) \to (\beta_1 z_1, \beta_2 z_2).$$

Hence we infer that G is a direct product $Z \times Z_l$ of an infinite cyclic group Z and a finite cyclic group Z_l. Thus we obtain the following

THEOREM 32. *If a Hopf surface $\boldsymbol{W}/G$ is non-elliptic, then G is a direct product $Z \times Z_l$ of an infinite cyclic group Z and a finite cyclic group Z_l of order l. Z is generated by a contradiction*

$$f : (z_1, z_2) \to (\alpha_1 z_1 + \lambda z_2{}^m, \alpha_2 z_2), \qquad (\alpha_1 - \alpha_2{}^m)\lambda = 0,$$

where m is a positive integer and $0 < |\alpha_1| \leqq |\alpha| < 1$. Z_l is generated by a linear transformation

$$e : (z_1, z_2) \to (\epsilon_1 z_1, \epsilon_2, z_2), \qquad (\epsilon_1 - \epsilon_2{}^m)\lambda = 0.$$

where ϵ_1 and ϵ_2 are primitive l-th roots of unity. Moreover the equation: $z_2 = 0$ defines a non-singular elliptic curve on the Hopf surface $\boldsymbol{W}/G$.

It is a corollary to this theorem that *every Hopf surface contains at least one non-singular elliptic curve.*

THEOREM 33. *Let S be an arbitrary surface. If a complex line bundle $F \in H^1(S, \mathcal{O}^*)$ has a finite order n, then there exists an unramified n-fold covering surface $\tilde{S}$ of S such that F induces on $\tilde{S}$ a trivial complex line bundle. If the Chern class $c(F)$ of a complex line bundle $F \in H^1(S, \mathcal{O}^*)$ has a finite order n, then there exists an unramified n-fold covering surface $\tilde{S}$ of S such that F induces on $\tilde{S}$ a complex line bundle of which the Chern class vanishes.*

Proof. Let $\{f_{jk}\}$ be a system of transition functions defining the complex line bundle F with respect to a finite covering $\{U_j\}$ of S by spherical neighborhoods U_j. Suppose that F [or $c(F)$] has a finite order n. Then there exists a 0-cochain $\{f_j\}$ composed of non-vanishing holomorphic [or differentiable] functions f_j defined respectively on U_j such that

$$f_{jk}{}^n = f_j / f_k, \qquad \text{on } U_j \cap U_k.$$

Let $g_{j\nu}=\exp(2\pi i\nu/n)f_j^{1/n}$, $\nu=0,1,\cdots,n-1$, be the n-th roots of f_j which are holomorphic [or differentiable] functions on U_j. We associate with $g_{j\nu}$ a copy $U_{j\nu}$ of U_j and, for any point $z\in U_j$, we denote by $(z)_{j\nu}$ the corresponding point on $U_{j\nu}$. Identifying $(z)_{j\nu}\in U_{j\nu}$ with $(z)_{k\mu}\in U_{k\mu}$ if and only if $g_{j\nu}=f_{jk}\cdot g_{k\mu}$, we form the union $\tilde{S}=\cup\, U_{j\nu}$. Obviously $\tilde{S}$ forms an unramified n-fold covering surface of S. We infer from the relation $g_{j\nu}=f_{jk}\cdot g_{k\mu}$ that F induces on $\tilde{S}$ a complex line bundle $\tilde{F}$ which is complex analytically [or differentiably] trivial, q. e. d.

THEOREM 34. *Let S be a surface with no non-constant meromorphic function. If $b_1(S)=1$, $b_2(S)=0$ and if S contains at least one curve, then S is a Hopf surface.*

To prove this theorem we first derive several lemmas. Let S be a surface satisfying the hypotheses of the theorem. For any sheaf Ξ over S we write $H^\nu(\Xi)$ in place of $H^\nu(S,\Xi)$. By Theorem 3 and the formula (13), S has the following numerical characters: $q=1$, $h^{1,0}=0$, $p_g=0$, $c_2=0$, $c_1^2=0$. Hence, by (14), the inclusion map $\boldsymbol{C}\to\mathcal{O}$ induces an isomorphism $H^1(\boldsymbol{C})\cong H^1(\mathcal{O})$. It follows that

$$\mathfrak{P}=H^1(\mathcal{O})/H^1(\boldsymbol{Z})\cong H^1(\boldsymbol{C})/H^1(\boldsymbol{Z})\cong\boldsymbol{C}^*. \tag{102}$$

We have the exact commutative diagram

$$\begin{tikzcd}
\cdots \arrow[r] & H^1(\boldsymbol{Z}) \arrow[r] \arrow[d, equal] & H^1(\boldsymbol{C}) \arrow[r] \arrow[d] & H^1(\boldsymbol{C}^*) \arrow[r] \arrow[d] & H^2(\boldsymbol{Z}) \arrow[r] \arrow[d, equal] & 0 \\
\cdots \arrow[r] & H^1(\boldsymbol{Z}) \arrow[r] & H^1(\mathcal{O}) \arrow[r] & H^1(\mathcal{O}^*) \arrow[r] & H^2(\boldsymbol{Z}) \arrow[r] & 0.
\end{tikzcd}$$

Hence we infer that the inclusion map $\boldsymbol{C}^*\to\mathcal{O}^*$ induces an isomorphism

$$H^1(\boldsymbol{C}^*)\cong H^1(\mathcal{O}^*). \tag{103}$$

This implies that any complex line bundle over S is represented by a 1-cocycle $\{c_{jk}\}$ on the nerve of a covering $\{U_j\}$ composed of constants $c_{jk}\in\boldsymbol{C}^*$. By hypothesis the second cohomology group $H^2(\boldsymbol{Z})$ is finite and therefore cc' vanishes for any elements $c,c'\in H^2(\boldsymbol{Z})$. Hence we infer from the Riemann-Roch-Hirzebruch theorem that, for any complex line bundle F over S,

$$\dim H^1(\mathcal{O}(F))=\dim H^0(\mathcal{O}(F))+\dim H^2(\mathcal{O}(F)). \tag{104}$$

LEMMA 9. *Any finite unramified covering surface $\tilde{S}$ of S satisfies the hypotheses of Theorem* 34.

Proof. We write $\tilde{S}=Q_m\cdots Q_2Q_1(S^*)$, where S^* is a surface free from exceptional curve and $Q_1,Q_2,\cdots,Q_m$ denote quadric transformations. It is clear that S^* has no non-constant meromorphic function. Hence, by Theorem 11, S^* is a K3 surface or a complex torus or $b_1(S^*)$ is equal to 1. Since $c_2(S)=0$, we have

$$c_2(S^*)+m=c_2(\tilde{S})=0.$$

This shows that S^* cannot be a K3 surface. If S^* were a complex torus, then $m=-c_2(S^*)$ would vanish and $\tilde{S}$ would be a complex torus. This contradicts that S is non-Kähler. Thus $b_1(S^*)$ is equal to 1 and therefore $b_2(S^*)=c_2(S^*)=-m$. Consequently m vanishes, $\tilde{S}$ coincides with S^*, and $b_2(\tilde{S})=b_2(S^*)=0$. Moreover $\tilde{S}$ contains a curve, e.g., the inverse image of a curve on S, q.e.d.

LEMMA 10. *S contains not more than two irreducible curves. Each irreducible curve on S is non-singular and elliptic.*

Proof. Let C denote an irreducible curve on S. Since

$$2\pi(C)-2=(C^2)+(KC)=0,$$

the virtual genus $\pi(C)$ of C is equal to 1 and C is either a non-singular elliptic curve or a rational curve with a singular point. We have the formula

$$\sum_{C}\pi(C)\leqq q-p_g+1=2$$

(see [2], Theorem 5.1). This proves that S contains not more than two irreducible curves.

For any positive integer n, there exists an n-fold unramified covering surface of S. To prove that C is not rational, we consider a 3-fold unramified covering surface $\tilde{S}$ of S and denote by ϖ the canonical projection of $\tilde{S}$ onto S. Suppose that C is rational and let

$$\varpi^{-1}(C)=C_1+\cdots+C_\nu+\cdots,$$

where the C_ν denote irreducible curves. Since, by Lemma 9, $b_2(\tilde{S})=0$, the curves C_ν are homologous to zero with respect to rational coefficients and therefore $(C_\lambda C_\nu)=0$. It follows that each curve C_ν is a connected component of $\varpi^{-1}(C)$. Consequently C_ν is an unramified covering space of C. Since $\pi(C)=1$, C has one and only one singular point which is either a cuspidal point or an ordinary double point. Hence we infer that ϖ maps C_ν biholomorphically onto C and that $\pi(C_\nu)=1$. This contradicts the formula

$$\sum_{\nu=1}^{3}\pi(C_\nu)\leqq q(\tilde{S})-p_g(\tilde{S})+1,$$

since, by Lemma 9, $p_g(\tilde{S}) = 0$ and $q(\tilde{S}) = 1$. Thus we see that C is a non-singular elliptic curve, q. e. d.

By a *multiplicative* holomorphic function on S we shall mean a multi-valued holomorphic function $w(z)$ defined on S such that the analytic continuation along any closed continuous curve γ transforms $w(z)$ into $\alpha w(z)$, where α is a constant depending on γ.

LEMMA 11. *For each irreducible curve C on S, there exists a multiplicative holomorphic function $w = w(z)$ whose divisor (w) is C.*

Proof. Let $\{U_j\}$ be a finite covering of S by small spherical neighborhoods U_j. The complex line bundle $[C]$ is represented by a 1-cocycle $\{c_{jk}\}$ on the nerve of the covering $\{U\}_j$ composed of constants $c_{ij} \in \mathbf{C}^*$. This implies the existence of a system $\{\phi_j\}$ of holomorphic functions ϕ_j on U_j such that

$$(\phi_j) = C, \qquad \phi_j = c_{jk}\phi_k \qquad\qquad \text{on } U_j \cap U_k.$$

Hence, setting

$$\eta = d\phi_j/\phi_j, \qquad\qquad \text{on each } U_j,$$

we obtain a d-closed meromorphic 1-form η on S whose polar cycle is C. We define

$$w = w(z) = \exp(\int^z \eta).$$

We have

$$w(z) = \phi_j(z) w(a)/\phi_j(a), \qquad\qquad \text{for } z, a \in U_j, a \notin C.$$

Hence w is a multiplicative holomorphic function whose divisor (w) is C, q. e. d.

For any 1-cycle γ on S with coefficients in $\mathbf{Z}$ which does not meet C, we define

$$\alpha(\gamma) = \exp(\int_\gamma w^{-1}\, dw).$$

It is clear that $\alpha(\gamma)$ is determined uniquely by the homology class of γ. Moreover, if γ is a closed continuous curve, the analytic continuation along γ induces the transformation: $w(z) \rightarrow \alpha(\gamma) w(z)$. The first Betti group of S is infinite cyclic. Let β denote a closed continuous curve representing a generator of the first Betti group of S. For any torsion cycle τ, $\alpha(\tau)$ is obviously a root of unity. We have

$$|\alpha(\beta)| \neq 1. \tag{105}$$

In fact, if $|\alpha(\beta)|=1$, then $|w(z)|$ would be a single-valued continuous function on S and attains its maximum. This contradicts that $w(z)$ is a non-constant holomorphic function.

LEMMA 12. *The curve C contains a 1-cycle γ which represents a homology class of infinite order on S.*

Proof. Let N be a tubular neighborhood of C in S. Suppose that all 1-cycles on C are torsion cycles on S. Then an arbitrary 1-cycle γ in N is a torsion cycle on S and therefore $\alpha(\gamma)^m=1$, where m denotes the order of the first torsion group of S. It follows that $w(z)^m$ is a single-valued homomorphic function on N whose divisor is mC. Hence, for any sufficiently small constant t, the equation: $w(z)^m=t$ defines a curve C_t in N. This contradicts Lemma 10, q. e. d.

Let $\mathfrak{U}$ denote the universal covering manifold of S. We identify the fundamental group $\pi_1(S)$ of S with the covering transformation group of $\mathfrak{U}$ with respect to S in a natural manner. Let Γ be the normal subgroup of $\pi_1(S)$ which corresponds to the infinite cyclic subgroup of $H_1(S, \boldsymbol{Z})$ generated by (the homology class of) β by means of the natural homomorphism: $\pi_1(S)\rightarrow H_1(S,\boldsymbol{Z})$. Obviously $\pi_1(S)/\Gamma$ is isomorphic to the first torsion group of S. Hence $\mathfrak{U}/\Gamma$ is a finite unramified covering surface of S.

i) *The case in which S contains two irreducible curves.* Let C_1 and C_2 be the irreducible curves on S which are, by Lemma 10, non-singular and elliptic. By Lemma 11, each curve C_ν, $\nu=1$ or 2, determines a multiplicative holomorphic function w_ν on S such that $(w_\nu)=C_\nu$. Note that $w_\nu^{-1}dw_\nu$ is a d-closed meromorphic 1-form on S. We let

$$\alpha_\nu(\gamma)=\exp\Big(\int_\gamma w_\nu^{-1}dw_\nu\Big).$$

Since C_2 does not meet C_1, the restriction of $w_1^{-1}dw_1$ to C_2 is a holomorphic 1-form on C_2. By Lemma 12, C_2 contains a 1-cycle γ which is homologous to $n\beta+\tau$, $n\neq 0$, where τ denotes a torsion cycle, and, by (105),

$$\Big|\exp\Big(\int_\gamma w_1^{-1}dw_1\Big)\Big|=|\alpha_1(\beta)|^n\neq 1.$$

Hence the restriction of $w_1^{-1}dw_1$ to C_2 does not vanish. Similarly the restriction of $w_2^{-1}dw_2$ to C_1 does not vanish. Since S contains no curves other than C_1 and C_2, it follows that the meromorphic 2-form $w_1^{-1}dw_1\wedge w_2^{-1}dw_2$ has the divisor

$$(w_1^{-1}dw_1\wedge w_2^{-1}dw_2)=-C_1-C_2.$$

Consequently *the 2-form* $dw_1 \wedge dw_2$ *does not vanish at each point on* S. By reversing the orientation of β if necessary, we may assume that

$$(106) \qquad |\alpha_\nu(\beta)| < 1, \qquad \text{for } \nu = 1, 2.$$

In fact, if, for instance, $\rho = -\log|\alpha_1(\beta)|$ and $\sigma = \log|\alpha_2(\beta)|$ were both positive, then $|w_1|^\sigma |w_2|^\rho$ would be a single-valued continuous function on S which attains its maximum at a point outside C_1 and C_2. This contradicts that $w_1{}^\sigma w_2{}^\rho$ is a non-constant holomorphic function in a neighborhood of the point.

The function w_ν can be considered as a single-valued holomorphic function $w_\nu(u)$ defined on the universal covering manifold $\mathfrak{U}$ of S, where u denotes a "variable" point on $\mathfrak{U}$. Since w_1 and w_2 do not vanish simultaneously,

$$\Phi\colon u \to \Phi(u) = (w_1(u), w_2(u))$$

is a locally biholomorphic map of $\mathfrak{U}$ into $\boldsymbol{W} = \boldsymbol{C}^2 - O$. We introduce a linear transformation

$$f\colon (w_1, w_2) \to (\alpha_1(\beta) w_1, \alpha_2(\beta) w_2).$$

By (106), f is a contraction. Since any element $\gamma \in \Gamma$ is homologous to a multiple of β, we have

$$\Phi(\gamma u) = f^n \Phi(u), \qquad \text{for } \gamma \in \Gamma,$$

where n is an integer depending on γ. Hence Φ induces a locally biholomorphic map of $\mathfrak{U}/\Gamma$ into the primary Hopf surface $\boldsymbol{W}/\{f\}$. This map is surjective, since $\mathfrak{U}/\Gamma$ is compact. It follows that $\Phi\colon \mathfrak{U} \to \boldsymbol{W}$ is surjective, while $\mathfrak{U}$ and $\boldsymbol{W}$ are both simply connected. Consequently Φ is a biholomorphic map of $\mathfrak{U}$ onto $\boldsymbol{W}$. Thus we conclude that S is a Hopf surface.

ii) *The case in which* S *contains one and only one irreducible curve.* Let C be the irreducible curve on S. By Lemma 10, C is non-singular and elliptic. We denote by K either the canonical bundle of S or a canonical divisor on S. Thus, for instance, we indicate by writing $K = -nC$ that $-nC$ is a canonical divisor on S. For any complex line bundle F over S we denote by F_C the restriction of F to the curve C. Moreover we denote by $\mathcal{O}(F + nC)$ the sheaf over S of germs of those meromorphic sections of F which are multiples of $-nC$ (compare [2], § 1).

LEMMA 13. *If the complex line bundle* $[C]_C$ *has a finite order* m*, then* S *is a Hopf surface and* $K = -(m+1)C$.

Proof. For any non-negative integer n, we have

$$(107)\qquad \dim H^2(\mathcal{O}(nC)) = \dim H^0(\mathcal{O}(K-nC)) = p_g(S) = 0.$$

Hence, using (104), we get

$$(108)\qquad \dim H^1(\mathcal{O}(nC)) = \dim H^0(\mathcal{O}(nC)) = 1.$$

Let $\iota_n\colon H^1(\mathcal{O}) \to H^1(\mathcal{O}(nC))$ and $\kappa_n\colon H^1(\mathcal{O}((n-1)C)) \to H^1(\mathcal{O}(nC))$ denote the homomorphisms induced by the inclusion maps: $\mathcal{O} \to \mathcal{O}(nC)$ and $\mathcal{O}((n-1)C) \to \mathcal{O}(nC)$, respectively. It is obvious that

$$(109)\qquad \kappa_n\iota_{n-1} = \iota_n, \qquad \text{for } n \geqq 1.$$

We have the exact sequence

$$0 \to \mathcal{O}((n-1)C) \to \mathcal{O}(nC) \to \mathcal{O}(n[C]_C) \to 0, \qquad \text{for } n \geqq 1.$$

Hence, using (107), we obtain the exact cohomology sequence

$$(110)\qquad \cdots \to H^1(\mathcal{O}((n-1)C)) \xrightarrow{\kappa_n} H^1(\mathcal{O}(nC)) \to H^1(C, \mathcal{O}(n[C]_C)) \to 0.$$

Since the order of $[C]_C$ is m, we have

$$(111)\qquad \dim H^1(C, \mathcal{O}(n[C]_C)) = \begin{cases} 0, & \text{for } n = 1, 2, \cdots, m-1, \\ 1, & \text{for } n = m. \end{cases}$$

Hence we infer from (109) and (110) that, if ι_{n-1} is surjective for an integer $n \leqq m-1$, then ι_n is surjective. Obviously ι_0 is surjective. Consequently $\iota_{m-1}\colon H^1(\mathcal{O}) \to H^1(\mathcal{O}((m-1)C))$ *is surjective.* It follows from (108), (110) and (111) that $\kappa_m H^1(\mathcal{O}((m-1)C))$ vanishes. Hence, by (109), $\iota_m H^1(\mathcal{O})$ *vanishes.*

By Lemma 11, there exists on S a multiplicative holomorphic function w with $(w) = C$. We take a finite covering $\{U_j\}$ of S by sufficiently small spherical neighborhoods U_j and denote by w_j a branch of w on U_j. Since the inclusion: $\mathbf{C} \subset \mathcal{O}$ induces an isomorphism: $H^1(\mathbf{C}) \cong H^1(\mathcal{O})$ and since $\dim H^1(\mathcal{O}) = 1$, there exists a 1-cocycle $\{c_{jk}\}$ on the nerve of $\{U_j\}$ composed of constants c_{jk} which generates $H^1(\mathcal{O})$. Since $\iota_m H^1(\mathcal{O})$ vanishes, $\{c_{jk}\}$ is cohomologous to zero in $\mathcal{O}(mC)$. This means the existence of a 0-cochain $\{h_j/w_j^m\}$ composed of sections h_j/w_j^m of $\mathcal{O}(mC)$ over the neighborhoods U_j such that

$$(112)\qquad c_{jk} = h_k/w_k^m - h_j/w_j^m, \qquad \text{on } U_j \cap U_k.$$

Hence, setting

$$\zeta = d(h_j/w_j{}^m), \qquad \text{on each } U_j,$$

we obtain a d-closed meromorphic 1-form ζ defined on S.

We shall prove that *the holomorphic function* h_j *is not constant on* C. The quotient $\alpha_{jk} = w_j/w_k$ is a constant. It follows from (112) that

$$h_j = \alpha_{jk}{}^m \cdot h_k, \qquad \text{on } C \cap U_j \cap U_k. \tag{113}$$

This shows that, if h_j is constant on C, then h_k is also constant on C. Suppose that, for each neighborhood U_j which meets C, the function h_j is equal to a constant a_j on C. If $a_j \neq 0$, then, by (113), $a_j = \alpha_{jk}{}^m a_k$ and therefore

$$x(z) = w_j{}^m/a_j = w_k{}^m/a_k = \cdot\cdot\cdot$$

is a holomorphic function on a neighborhood N of C whose divisor (x) is mC. Hence, for any sufficiently small constant t, the equation: $x(z) = t$ defines a curve C_t in N. This contradicts Lemma 10. Thus the constants a_j vanish and therefore the quotients h_j/w_j are holomorphic functions. Hence the formula (112) implies that $\{c_{jk}\}$ is cohomologous to zero in $\mathcal{O}((m-1)C)$. This contradicts that $\iota_{m-1}\colon H^1(\mathcal{O}) \to H^1(\mathcal{O}((m-1)C))$ is surjective.

We define

$$w_1 = w_1(z) = w(z)^m \int^z \zeta.$$

We have

$$w_1 = \alpha_j h_j + \lambda_j w^m, \qquad \text{on } U_j,$$

where α_j and λ_j are constants depending on the branch of w_1 and $\alpha_j \neq 0$. Hence w_1 is a multi-valued holomorphic function on S. The analytic continuation along a closed continuous curve γ induces the transformation:

$$w_1 \to \alpha(\gamma)^m (w_1 + \mu(\gamma) w^m), \tag{114}$$

where

$$\alpha(\gamma) = \exp\Big(\int_\gamma w^{-1}\, dw\Big), \quad \mu(\gamma) = \int_\gamma \zeta.$$

It follows that, on the elliptic curve C, w_1 is a non-constant multiplicative holomorphic function. Hence w_1 does not vanish at each point of C. We have

$$w^{-1}\, dw \wedge \zeta = w_j{}^{-m-1}\, dw_j \wedge dh_j.$$

Since S contains no curves other than C and since h_j is not constant on C, we infer that the meromorphic 2-form $w^{-1}\, dw \wedge \zeta$ has the divisor

$$(w^{-1}\, dw \wedge \zeta) = -(m+1)C.$$

This proves that $K = -(m+1)C$ and that *the 2-form*

$$dw_1 \wedge dw = w^m \, dw \wedge \zeta$$

does not vanish at each point of S. We consider w_1 and w, respectively, as single-valued holomorphic functions $w_1(u)$ and $w(u)$ defined on the universal covering manifold $\mathfrak{U}$ of S. Since w_1 and w do not vanish simultaneously,

$$\Phi \colon u \rightarrow \Phi(u) = (w_1(u), w(u))$$

is a locally biholomorphic map of $\mathfrak{U}$ into $\mathbf{W}$. We may assume that $|\alpha(\beta)| < 1$. We introduce a contraction

$$f \colon (w_1, w) \rightarrow (\alpha^m w_1 + \lambda w^m, \alpha w), \qquad \alpha = \alpha(\beta), \lambda = \alpha^m \mu(\beta).$$

We infer from (114) that

$$\Phi(\gamma u) = f^n \Phi(u), \qquad \text{for } \gamma \in \Gamma,$$

where n is an integer depending on γ. Therefore Φ induces a locally biholomorphic map of $\mathfrak{U}/\Gamma$ into $\mathbf{W}/\{f\}$. Hence we conclude in the same manner as in i) above that S is a Hopf surface, q. e. d.

LEMMA 14. *The divisor $-C$ is not canonical.*

Proof. Let $\mathcal{O}_C$ and $\mathcal{O}^*_C$ denote respectively the restrictions of $\mathcal{O}$ and $\mathcal{O}^*$ to C. Let r denote the restriction map to C. We have the exact commutative diagram

$$\begin{tikzcd}
 & \vdots \arrow[d] & & & \\
\cdots \arrow[r] & H^1(\mathcal{O}) \arrow[r, "e"] \arrow[d, "r"] & H^1(\mathcal{O}^*) \arrow[r] \arrow[d, "r"] & H^2(\mathbf{Z}) \arrow[r] & 0 \\
\cdots \arrow[r] & H^1(C, \mathcal{O}_C) \arrow[r, "e"] \arrow[d] & H^1(C, \mathcal{O}^*_C) \arrow[r] & \cdots & \\
 & H^2(\mathcal{O}(-C)) \arrow[d] & & & \\
 & 0 & & &
\end{tikzcd} \tag{115}$$

Suppose that $K = -C$. Then we have

$$\dim H^2(\mathcal{O}(-C)) = \dim H^0(\mathcal{O}) = 1.$$

Hence we infer from (115) that $rH^1(\mathcal{O}) = 0$. Let n denote the order of the finite group $H^2(\boldsymbol{Z})$. Take any complex line bundle $F\in H^1(\boldsymbol{\mathcal{O}}^*)$. Since $c(nF)=nc(F)=0$, there exists an element $\xi\in H^1(\boldsymbol{\mathcal{O}})$ such that $nF=e\xi$. Hence, using (115), we get

$$nrF=rnF=re\xi=er\xi=0.$$

It follows that $[C]_C=r[C]$ has a finite order m. Therefore, by Lemma 13, $-(m+1)C$ is a canonical divisor. Combined with $K=-C$, this proves that $[mC]=0$. This is a contradiction, q. e. d.

Since C is a non-singular elliptic curve, the canonical bundle $K_C+[C]_C$ is trivial. Hence, for any positive integer n, we have the exact sequence

$$(116)\qquad 0\to\boldsymbol{\mathcal{O}}(nK+(n-1)C)\to\boldsymbol{\mathcal{O}}(nK+nC)\to\boldsymbol{\mathcal{O}}_C\to 0.$$

With the aid of (104) and of the duality theorem, we infer from the exact cohomology sequence derived from (116) that either $|nK+nC|$ or $|-(n-1)(K+C)|$ is non-empty. Since S contains no curves other than C, it follows that either $|nK+nC|=|k_nC|$ or $|-(n-1)(K+C)|=|h_nC|$, where k_n and h_n denote non-negative integers. Hence we infer that either $K=-l[C]$ or $2K+2[C]=0$, where l is an integer $\geqq 2$. In fact, if $|-K-C|=|h_2C|$, then we have $K=-l[C]$, $l=h_2+1$. If $|2K+2C|=|k_2C|$ and if $|3K+3C|=|k_3C|$, then we have $K=-l[C]$, where $l=k_2-k_3+1$. If $|2K+2C|=|k_2C|$ and if $|-2K-2C|=|h_3C|$, then we have $2K+2[C]=0$. Since $p_g=0$, the integer l is positive. Moreover, by Lemma 14, l is not equal to 1.

If $K=-l[C]$, $l\geqq 2$, then, letting $n=l-1$, we obtain

$$n[C]_C=[lC-C]_C=-K_C-[C]_C=0.$$

Hence, by Lemma 13, S is a Hopf surface.

If $2K+2[C]=0$, then, by Theorem 33, there exists a 2-fold unramified covering surface $\tilde{S}$ of S such that $\tilde{K}+[\tilde{C}]=0$, where $\tilde{K}$ is the canonical bundle of $\tilde{S}$ and $\tilde{C}$ is the inverse image of C on $\tilde{S}$ with respect to the covering map. By Lemma 9, $\tilde{S}$ satisfies the hypotheses of Theorem 34. Moreover $\tilde{S}$ contains no curves other than the components of $\tilde{C}$. Hence, by Lemma 14, $\tilde{C}$ cannot be irreducible. Thus $\tilde{C}$ is composed of two irreducible curves. Therefore, by the result of i), $\tilde{S}$ is a Hopf surface. Consequently S is a Hopf surface. This completes the proof of Theorem 34.

THEOREM 35. *An elliptic surface S of class* VII_0 *is a Hopf surface if and only if the* 12-*genus P_{12} of S vanishes.*

Proof. In view of Theorem 28, it suffices to show that $\kappa(S)$ is negative if and only if P_{12} vanishes. We recall that

$$K = -2C_u + \sum_{\nu=1}^{r} (m_\nu - 1) P_\nu$$

is a canonical divisor of S and that

$$\kappa = \kappa(S) = -2 + \sum_{\nu=1}^{r} (1 - 1/m_\nu).$$

$m_\nu P_\nu$ is linearly equivalent to C_u. Hence, for any positive integer m which is divisible by m_ν, $\nu = 1, 2, \cdots, r$, we have

$$| mK | = | m\kappa C_u |.$$

We infer therefore that, if $\kappa < 0$, then all pluri-genera of S vanish.

Assume that P_{12} vanishes. We determine integers k_ν and s_ν such that

$$m_\nu k_\nu = 12 + s_\nu, \qquad 0 \leqq s_\nu < m_\nu.$$

Note that $k_\nu \leqq 6$ and $k_\nu \neq 5$. We have

$$| 12K | = | (-24 + \sum_{\nu=1}^{r} (12 - k_\nu)) C_u + \sum_{\nu=1}^{r} s_\nu P_\nu |,$$

while, by hypothesis, $| 12K |$ is empty. Hence we get

$$\sum_{\nu=1}^{r} (12 - k_\nu) < 24. \tag{117}$$

This shows that $r \leqq 3$. If $r \leqq 2$, then κ is obviously negative. Therefore it suffices to consider the case in which $r = 3$. Assume that $k_1 \geqq k_2 \geqq k_3$. The inequality (117) says that

$$k_1 + k_2 + k_3 > 12.$$

Hence we infer that either a) $k_1 = k_2 = 6$ or b) $k_1 = 6$, $k_2 = 4$, $k_3 \geqq 3$. In the case a), we get $m_1 = m_2 = 6$. In the case b), we get $m_1 = 2$, $m_2 = 3$, $m_3 \leqq 5$. Consequently we obtain

$$\kappa = 1 - \sum_{\nu=1}^{3} 1/m_\nu < 0,$$

q. e. d.

THEOREM 36. *Any deformation of any Hopf surface is a Hopf surface.*

Proof. In view of Theorem 30, it suffices to prove that any deformation of a primary Hopf surface is a (primary) Hopf surface. Any primary Hopf surface is a quotient surface $S=W/\{f\}$, where f is a contraction of the form (94), i. e.,

$$f\colon (z_1,z_2)\to(\alpha_1 z_1+\lambda z_2{}^m,\alpha_2 z_2).$$

First we show that any sufficiently small deformation S_t of S is a Hopf surface. In the case in which $\alpha_1\neq\alpha_2$, it follows from (99) that the curve C defined by the equation: $z_2=0$ is a stable submanifold of S (see [4], Theorem 1). This means that any sufficiently small deformation S_t of S contains a deformation C_t of C. Moreover it follows from (97) that the 12-genus $P_{12}(S_t)$ vanishes, since $P_{12}(S_t)$ is an upper semi-continuous function of t. Hence, by Theorems 34 and 35, S_t is a Hopf surface.

In the case in which $\alpha_1=\alpha_2$, f is a linear transformation. For any linear transformation g of $\boldsymbol{C}^2$ which is in a small neighborhood N of f, we define a primary Hopf surface $S_g=W/\{g\}$. The collection of the Hopf surfaces S_g, $g\in N$, form a complete complex analytic family (see Kodaira and Spencer [7], Theorem 15.4 and [8]). Hence any sufficiently small deformation of $S=S_f$ is a Hopf surface.

Let $\{S_t \mid t\in M\}$ be a complex analytic family of deformations S_t of a primary Hopf surface S, where M denotes a connected complex manifold. Let N be the subset of M determined by the condition that S_t is a Hopf surface if and only if $t\in N$. Our purpose is to show that N coincides with M. By the above result, N is a non-empty open subset of M. Hence it suffices for our purpose to prove that N is closed. Take a point b on the boundary of N. Let K_t denote the canonical bundle of S_t. By (97), $\dim|-K_t|$ is non-negative for $t\in N$. Hence, by the principle of upper semi-continuity, $\dim|-K_b|$ is also non-negative, while, since $p_g(S_b)=p_g(S)=0$, K_b is non-trivial. Therefore the anti-canonical system $|-K_b|$ contains a positive divisor and $P_{12}(S_b)$ vanishes. Consequently, by Theorems 34 and 35, S_b is a Hopf surface and b belongs to N, q. e. d.

THEOREM 37. *Any deformation of any elliptic surface of class* VII_0 *is either an elliptic surface or a Hopf surface.*

Proof. Consider a complex analytic family $\{S_t \mid t\in M\}$ of surfaces S_t, $t\in M$, such that, for a point $o\in M$, S_o is an elliptic surface of class VII_0. It is obvious that $b_1(S_t)=1$ and $b_2(S_t)=0$. It follows that S_t contains no exceptional curve and $p_g(S_t)=0$. Hence S_t belongs to the class VII_0. If $P_{12}(S_o)$ vanishes, then, by Theorem 35, S_o is a Hopf surface and therefore, by Theorem 36, any deformation S_t of S_o is a Hopf surface.

Assume that $P_{12}(S_o)$ is positive. Then $\kappa(S_o)$ is non-negative, as was shown in the proof of Theorem 35, and therefore S_o has at least three singular fibres. Hence S_o belongs to a complete complex analytic family $\mathfrak{F}$ as described in Theorem 29. Since all members of $\mathfrak{F}$ have the same pluri-genera, it follows that any sufficiently small deformation S_t of S_o is an elliptic surface with $P_{12}(S_t) = P_{12}(S_o) \geqq 1$.

Let N denote the subset of M determined by the condition that S_t is an elliptic surface with $P_{12}(S_t) \geqq 1$ if and only if $t \in N$. By the above result, N is an open subset of M. For our purpose it suffices therefore to show that N is closed. Take a point b on the boundary of N. By the principle of upper semi-continuity, $P_{12}(S_b)$ is positive. Hence the multi-canonical system $|12K_b|$ is non-empty, where K_b denotes the canonical bundle of S_b. If $12K_b$ is trivial, then, by Theorem 33, there exists a finite unramified covering surface $\tilde{S}_b$ of S_b whose canonical bundle is trivial. By Theorem 19, $\tilde{S}_b$ is a complex torus, a K3 surface or an elliptic surface, while, since $b_1(S_b) = 1$ and $b_2(S_b) = 0$, $\tilde{S}_b$ is non-Kähler and $c_2(\tilde{S}_b) = 0$. Hence $\tilde{S}_b$ is elliptic and consequently S_b is elliptic. Thus b belongs to N. If $12K_b$ is non-trivial, then $|12K_b|$ contains a positive divisor. If S_b were non-elliptic, then, by Theorems 34 and 36, S_o would be a Hopf surface. This contradicts Theorem 35. Consequently S_b is elliptic and b belongs to N, q. e. d.

11. Surfaces with volume-preserving complex structure.

By a *volume-preserving complex structure* we shall mean a pseudo-group structure corresponding to the pseudo-group of volume-preserving local complex analytic homeomorphisms of the space of several complex variables. In this section we shall examine the application of our results to surfaces with volume-preserving complex structures.

Let S be a surface with a volume-preserving complex structure. The volume-preserving complex structure of S is defined by a system $\{(z_j^1, z_j^2)\}$ of local complex coordinates (z_j^1, z_j^2) such that

$$dz_j^1 \wedge dz_j^2 = \kappa_{jk}\, dz_k^1 \wedge dz_k^2, \qquad |\kappa_{jk}| = 1.$$

Let U_j denote the domain of the local coordinate (z_j^1, z_j^2). The canonical bundle K of S is defined by the 1-cocycle $\{\kappa_{jk}^{-1}\}$ on the nerve of the covering $\{U_j\}$ composed of the *constants* κ_{jk}^{-1}. For any sheaf Ξ over S we write $H^\nu(\Xi)$ for $H^\nu(S, \Xi)$. We have the exact commutative diagram

$$\begin{tikzcd}
\cdots \arrow[r] & H^1(\boldsymbol{Z}) \arrow[r] \arrow[d, equal] & H^1(\boldsymbol{C}) \arrow[r] \arrow[d] & H^1(\boldsymbol{C}^*) \arrow[r] \arrow[d] & H^2(\boldsymbol{Z}) \arrow[r] \arrow[d, equal] & H^2(\boldsymbol{C}) \arrow[r] \arrow[d] & \cdots \\
\cdots \arrow[r] & H^1(\boldsymbol{Z}) \arrow[r] & H^1(\mathcal{O}) \arrow[r] & H^1(\mathcal{O}^*) \arrow[r] & H^2(\boldsymbol{Z}) \arrow[r] & H^2(\mathcal{O}) \arrow[r] & \cdots
\end{tikzcd} \tag{118}$$

Since $K \in H^1(\mathcal{O}^*)$ is the image of the cohomology class $\in H^1(\boldsymbol{C}^*)$ of the 1-cocycle $\{\kappa_{jk}^{-1}\}$, we infer from (118) that the image of $c(K) \in H^2(\boldsymbol{Z})$ in $H^2(\boldsymbol{C})$ vanishes. It follows that $c_1 = -c(K)$ *is a torsion element of* $H^2(\boldsymbol{Z})$.

THEOREM 38. *Any surface S with a volume-preserving complex structure has a finite unramified normal covering surface of which the canonical bundle is trivial.*

Proof. It suffices to prove that S has a finite unramified covering surface with a trivial canonical bundle. The order of the first Chern class c_1 of S is finite. Hence, by Theorem 33, there is a finite unramified covering surface $\tilde{S}$ of S of which the first Chern class vanishes. Obviously $\tilde{S}$ has a volume-preserving complex structure. Therefore, in view of Theorem 33, it suffices for our purpose to prove *the finiteness of the order of the canonical bundle* $K \in H^1(S, \mathcal{O}^*)$ *of S under the assumption that* $c_1 = 0$.

For convenience we consider the group $H^1(S, \mathcal{O}^*)$ of complex line bundles as a *multiplicative group*. Thus, for instance, we denote the trivial complex line bundle by 1. We suppose that K^n is non-trivial for any integer $n \neq 0$ and derive a contradiction. We have

$$\dim H^0(S, \mathcal{O}(K^n)) = 0, \qquad \text{for } n \neq 0. \tag{119}$$

In fact, if ϕ is a holomorphic section of K^n, then the fibre coordinates ϕ_j of ϕ satisfy $\phi_j = \kappa_{jk}^{-n}\phi_k$ and therefore $|\phi_j| = |\phi_k| = \cdots$ is a continuous function on S which attains its maximum. It follows that the ϕ_j are constants, while K^n is non-trivial. Hence ϕ vanishes.

We shall examine the numerical characters p_g, q, $h^{1,0}$, b_1, and b_2 of S. The irregularity q is positive. In fact, since, by hypotheses, $c_1 = 0$, the vanishing of q would imply that K is trivial (see § 3). By (119), p_g is equal to zero. Hence, by (13), q is equal to 1 and therefore b_1 is equal to 2 or 1.

Suppose that $b_1 = 2$. Then the Albanese variety $\mathfrak{A}$ attached to S is an elliptic curve and there is a holomorphic map Φ of S onto $\mathfrak{A}$ defined by (15). For any general point $u \in \mathfrak{A}$, the inverse image $C_u = \Phi^{-1}(u)$ is a non-singular curve. Moreover, since $c_1 = 0$, (KC_u) vanishes and therefore C_u is elliptic. Thus S is an elliptic surface and consequently, by Theorem 12,

$$K = \Phi^*(-\mathfrak{f}) + [\sum (m_\nu - 1)P_\nu],$$

where $\mathfrak{f}$ is a complex line bundle over $\mathfrak{A}$ with $c(\mathfrak{f}) = 0$. It follows that the effective divisor $\sum (m_\nu - 1)P_\nu$ is homologous to zero, while, by Theorem 10, S is algebraic. Hence $\sum (m_\nu - 1)P_\nu$ vanishes. Moreover the Euler number

$c_2 = 12(p_g - q + 1)$ of S is zero. Consequently S has no singular fibre and the fibre C_u is complex analytically homeomorphic to an elliptic curve C which is independent of u. Employing the notation of § 9, we denote a point on C by $[\zeta]$. Choose a covering $\{E_j\}$ of $\mathfrak{A}$ by small disks E_j. Then $\Phi^{-1}(E_j)$ is a cartesian product $E_j \times C$ and $(u, [\zeta_j]) \in \Phi^{-1}(E_j)$ is identical with $(u, [\zeta_k]) \in \Phi^{-1}(E_k)$ if and only if

$$[\zeta_j] = [\epsilon_{jk}\zeta_k + \eta_{jk}(u)],$$

where $\eta_{jk}(u)$ is a holomorphic function on $E_j \cap E_k$ and ϵ_{jk} is ± 1, $\pm i$ or $\frac{1}{2}(\pm 1 \pm i\sqrt{3})$. It follows that

$$du \wedge d\zeta_j = \epsilon_{jk}\, du \wedge d\zeta_k,$$

where du is a holomorphic 1-form on $\mathfrak{A}$. Hence K is defined by the 1-cocycle $\{\epsilon_{jk}\}$ and consequently $K^{12} = 1$.

Thus b_1 is equal to 1. We then infer from (13) and Theorem 3 that $b_2 = 0$ and $h^{1,0} = 0$. It follows that S belongs to the class VII_0.

Let $\mathfrak{P}$ denote the Picard variety attached to S. Since $h^{1,0} = 0$, the formula (14) shows that the inclusion $\boldsymbol{C} \subset \boldsymbol{\mathfrak{O}}$ induces an isomorphism: $H^1(\boldsymbol{C}) \cong H^1(\boldsymbol{\mathfrak{O}})$, while $b_2 = p_g = 0$. Hence we obtain from (118) the isomorphism

$$\mathfrak{P} \cong H^1(\boldsymbol{C})/H^1(\boldsymbol{Z}) \cong \boldsymbol{C}/\boldsymbol{Z} = \boldsymbol{C}^*. \tag{120}$$

Let $\{m_{jk}\}$ denote a 1-cocycle on the nerve of the covering $\{U_j\}$ which represents the generator of $H^1(\boldsymbol{Z})$. We then infer from (120) that, for any complex number $F \in \boldsymbol{C}^*$, the corresponding complex line bundle in $\mathfrak{P} \cong \boldsymbol{C}^*$ is defined by the 1-cocycle $\{F_{jk}\}$, $F_{jk} = F^{m_{jk}}$. For the sake of simplicity, we use the same symbol F to denote the complex line bundle defined by $\{F^{m_{jk}}\}$. Thus *we identify* $\mathfrak{P}$ *with* $\boldsymbol{C}^*$. The canonical bundle $K \in \mathfrak{P} = \boldsymbol{C}^*$ is defined by the 1-cocycle $\{\kappa_{jk}^{-1}\}$ composed of constants κ_{jk}^{-1} with $|\kappa_{jk}| = 1$. It follows that the absolute value of the complex number K is equal to 1.

S contains no curve. In fact, if S is elliptic, then, by Theorem 35, S is a Hopf surface, since, by (119), $P_{12}(S)$ vanishes. If S is non-elliptic and contains a curve, then, by Theorem 34, S is a Hopf surface. Thus, if S contains a curve, S is a Hopf surface and therefore, by Theorem 30, there exists a finite unramified covering surface $\tilde{S}$ of S which is a primary Hopf surface. Let n denote the number of sheets of the covering surface $\tilde{S}$ and let $\tilde{K}$ be the canonical bundle of $\tilde{S}$. By (97) the anti-canonical system $|\tilde{K}^{-1}|$ is non-empty. Hence we infer that $|K^{-n}|$ is non-empty. This contradicts (119).

Take a complex line bundle $F \in \mathfrak{P}$ and denote by $\Omega^{\nu}(F)$ the sheaf over S of germs of holomorphic ν-forms with coefficients in F. Since F is defined by a 1-cocycle $\{F^{m_{jk}}\}$ composed of *constants* $F^{m_{jk}}$, the exterior differentiation $d\colon \Omega^{\nu}(F) \to \Omega^{\nu+1}(F)$ is well-defined and the sequence

$$0 \to \boldsymbol{C}(F) \to \boldsymbol{\mathcal{O}}(F) \xrightarrow{d} \boldsymbol{\Omega}^1(F) \xrightarrow{d} \boldsymbol{\Omega}^2(F) \to 0$$

is exact, where $\boldsymbol{C}(F)$ denotes the sheaf over S of germs of *constant sections* of F. Hence we obtain the following exact sequences:

$$0 \to \boldsymbol{C}(F) \to \boldsymbol{\mathcal{O}}(F) \xrightarrow{d} d\boldsymbol{\mathcal{O}}(F) \to 0, \tag{121}$$

$$0 \to d\boldsymbol{\mathcal{O}}(F) \to \Omega^1(F) \xrightarrow{d} \Omega^2(F) \to 0. \tag{122}$$

Since S contains no curve, we have

$$\dim H^0(S, \boldsymbol{\mathcal{O}}(F)) = 0, \qquad \text{for } F \neq 1, \tag{123}$$

while we obtain from the Riemann-Roch-Hirzebruch theorem the equality

$$\dim H^1(\boldsymbol{\mathcal{O}}(F)) = \dim H^0(\boldsymbol{\mathcal{O}}(F)) + \dim H^2(\boldsymbol{\mathcal{O}}(F)). \tag{124}$$

Hence, using the duality theorem, we get

$$\dim H^{\nu}(S, \boldsymbol{\mathcal{O}}(F)) = 0, \qquad \text{for } \nu = 0, 1, 2, F \neq 1, \neq K. \tag{125}$$

Combining this with the exact cohomology sequence derived from (121), we obtain

$$H^{\nu}(S, d\boldsymbol{\mathcal{O}}(F)) \cong H^{\nu+1}(S, \boldsymbol{C}(F)), \quad \text{for } \nu = 0, 1, 2, F \neq 1, \neq K. \tag{126}$$

We have the isomorphism: $\Omega^2(F) \cong \boldsymbol{\mathcal{O}}(K \cdot F)$. Hence, using (125), we infer from the exact cohomology sequence derived from (122) that

$$H^{\nu}(S, d\boldsymbol{\mathcal{O}}(F)) \cong H^{\nu}(S, \Omega^1(F)), \qquad \text{for } \nu = 0, 1, 2, F \neq 1, \neq K^{-1}.$$

Combining this with (126), we get

$$H^{\nu}(S, \Omega^1(F)) \cong H^{\nu+1}(S, \boldsymbol{C}(F)), \qquad \text{for } F \neq 1, \neq K^{\pm 1}. \tag{127}$$

Now we substitute K^{-1} for F in the exact sequence (122). Let Θ denote the sheaf over S of germs of holomorphic vector fields. $\Omega^1(K^{-1})$ is the sheaf over S of germs of holomorphic tensor fields τ of the type $\tau_{\lambda}{}^{\alpha\beta}$, $\tau_{\lambda}{}^{\alpha\beta} = -\tau_{\lambda}{}^{\beta\alpha}$. The tensor $\tau_{\lambda}{}^{\alpha\beta}$ has only two independent components $\tau_1{}^{21}$ and $\tau_2{}^{12}$. Hence the map

$$\tau \to \theta = \theta^1 \partial_1 + \theta^2 \partial_2, \qquad \theta^1 = \tau_2{}^{12}, \theta^2 = \tau_1{}^{21},$$

defines an isomorphism: $\Omega^1(K^{-1}) \cong \Theta$, where the symbol ∂_α stands for $\partial/\partial z_j{}^\alpha$. Similarly $\Omega^2(K^{-1})$ is the sheaf of germs of holomorphic tensor fields σ of the type $\sigma_{\lambda\mu}{}^{\alpha\beta}$, $\sigma_{\lambda\mu}{}^{\alpha\beta} = -\sigma_{\lambda\mu}{}^{\beta\alpha} = -\sigma_{\mu\lambda}{}^{\alpha\beta}$, and the map: $\sigma \to \sigma_{12}{}^{12}$ defines an isomorphism: $\Omega^2(K^{-1}) \cong \mathcal{O}$. We have

$$(d\tau)_{12}{}^{12} = \partial_1 \tau_2{}^{12} - \partial_2 \tau_1{}^{12} = \partial_1 \theta^1 + \partial_2 \theta^2 = \delta\theta.$$

Note that *the divergent* $\delta\theta = \partial_1\theta^1 + \partial_2\theta^2$ *is invariant under any volume-preserving transformation of local complex coordinates.* Thus we obtain from (122) the exact sequence

$$0 \to d\mathcal{O}(K^{-1}) \to \Theta \xrightarrow{\delta} \mathcal{O} \to 0.$$

Using (126), we derive from this the exact sequence

$$(128) \qquad \begin{aligned} 0 &\to H^1(\mathbf{C}(K^{-1})) \to H^0(\Theta) \xrightarrow{\delta} H^0(\mathcal{O}) \to H^2(\mathbf{C}(K^{-1})) \\ &\to H^1(\Theta) \xrightarrow{\delta} H^1(\mathcal{O}) \to H^3(\mathbf{C}(K^{-1})) \to H^2(\Theta) \to 0. \end{aligned}$$

We have

$$(129) \qquad \dim H^0(\Theta) \cdot \dim H^0(\Omega^1(F)) = 0, \qquad \text{for } F \neq K^{-1}.$$

In fact, any element ϕ of $H^0(\Omega^1(F))$ is a 0-cochain $\{\phi_j\}$ composed of holomorphic 1-forms $\phi_j = \sum_\alpha \phi_{j\alpha} dz_j{}^\alpha$ such that $\phi_j = F^{m_{jk}}\phi_k$. If $H^0(\Theta)$ contains $\theta \neq 0$, then $\phi\theta = \{\sum_\alpha \phi_{j\alpha}\theta^\alpha\}$ vanishes, since it belongs to $H^0(\mathcal{O}(F))$ and since ϕ vanishes if $F = 1$. Hence there exists a meromorphic function f_j on each neighborhood U_j such that $\phi_{j1} = f_j\theta^2$, $\phi_{j2} = -f_j\theta^1$. It is easy to verify that $f_j = (KF)^{m_{jk}} \cdot f_k$. Thus $f = \{f_j\}$ is a meromorphic section of the complex line bundle KF, while S contains no curve and $KF \neq 1$. Consequently f vanishes. Thus we obtain (129). In a similar manner we get

$$(130) \qquad \dim H^0(\Theta) \leqq 1.$$

In fact, if $\theta, \sigma \in H^0(\Theta)$, $\theta \neq 0$, then the tensor $\theta^\alpha\sigma^\beta - \theta^\beta\sigma^\alpha$ vanishes, since it belongs to $H^0(\mathcal{O}(K^{-1}))$. Hence σ^α is equal to $f\theta^\alpha$ and the meromorphic function f is reduced to a constant. Combining (129) with the duality theorem, we get

$$(131) \qquad \dim H^0(\Theta) \cdot \dim H^2(\Theta) = 0.$$

Moreover, using the Riemann-Roch-Hirzebruch theorem, we obtain

$$\dim H^1(\Theta) = \dim H^0(\Theta) + \dim H^2(\Theta). \tag{132}$$

We have the *duality theorem*:

$$\dim H^{\nu}(S, \boldsymbol{C}(F^{-1})) = \dim H^{4-\nu}(S, \boldsymbol{C}(F)).$$

Obviously $\dim H^{\nu}(S, \boldsymbol{C}(F))$ is invariant under the conjugation: $F \rightarrow \bar{F}$ and $K^{-1} = \bar{K}$. Consequently we get

$$\dim H^1(\boldsymbol{C}(K^{-1})) = \dim H^3(\boldsymbol{C}(K)) = \dim H^1(\boldsymbol{C}(K)). \tag{133}$$

Now, using (130), (131), (132) and (133), we infer from the exact sequence (128) that $\dim H^1(\boldsymbol{C}(K)) \leqq 1$ and that

$$\dim H^0(\Theta) = \dim H^1(\Theta) = 1, \qquad \dim H^2(\Theta) = 0. \tag{134}$$

i) First we consider the case in which $\dim H^1(\boldsymbol{C}(K)) = 0$. We infer from (134) that there exists a complex analytic family $\mathcal{S} = \{S_t\}$ of surfaces S_t depending on a complex parameter t, $|t| < \epsilon$, $\epsilon > 0$, and satisfying the boundary condition: $S_0 = S$ such that the infinimal deformation $\partial S_t/\partial t$ does not vanish at $t = 0$ (see Kodaira, Nirenberg and Spencer [6]). Let $\mathcal{P}_t$ denote the Picard variety attached to S_t. Since $b_1(S_t) = q(S_t) = 1$, we have the isomorphism

$$\mathcal{P}_t \cong H^1(S_t, \boldsymbol{C})/H^1(S_t, \boldsymbol{Z}) \cong \boldsymbol{C}/\boldsymbol{Z} = \boldsymbol{C}^*.$$

By means of this isomorphism we identify $\mathcal{P}_t$ with $\boldsymbol{C}^*$. Since $c_1(S_t) = c_1 = 0$, the canonical bundle of S_t belongs to $\mathcal{P}_t \cong C^*$, i.e., the canonical bundle of S_t is defined by a 1-cocycle $\{K_t^{m_{jk}}\}$, $K_t \in \boldsymbol{C}^*$. We infer readily that K_t depends holomorphically on t. Moreover we have local complex coordinates (z_j^1, z_j^2, t) on the family $\mathcal{S}$ such that, for each fixed t,

$$dz_j^1 \wedge dz_j^2 = K_t^{-m_{jk}} dz_k^1 \wedge dz_k^2. \tag{135}$$

Let $z_j^{\alpha} = f_{jk}^{\alpha}(z_k, t)$. The infinitesimal deformation $\partial S_t/\partial t$ is the cohomology class of the 1-cocycle $\{\theta_{jk}(t)\}$ composed of local holomorphic vector fields

$$\theta_{jk}(t) = \sum \theta_{jk}^{\alpha} \partial/\partial z_j^{\alpha}, \qquad \theta_{jk}^{\alpha} = \partial f_{jk}^{\alpha}(z_k, t)/\partial t.$$

From (135) we get

$$d\theta_{jk}^1 \wedge dz_j^2 - d\theta_{jk}^2 \wedge dz_j^1 = - m_{jk} (\partial K_t/\partial t) K_t^{-m_{kj}-1} dz_k^1 \wedge dz_k^2.$$

Hence we obtain

$$\delta\theta_{jk}(t) = - m_{jk} K_t^{-1} \partial K_t/\partial t. \tag{136}$$

Now we set $t=0$. Since $\dim H^3(\boldsymbol{C}(K^{-1}))=\dim H^1(\boldsymbol{C}(K))=0$, we infer from (128) and (134) that $\delta: H^1(\Theta)\to H^1(\mathcal{O})$ is an isomorphism. Moreover the 1-cocycle $\{\theta_{jk}(0)\}$ is not cohomologous to zero, since it represents the infinitesimal deformation $(\partial S_t/\partial t)_{t=0}\neq 0$. Therefore (136) proves that $\partial K_t/\partial t$ does not vanish at $t=0$. Thus K_t is a non-constant holomorphic function of t, while the absolute value of K_0 is equal to 1. Hence there exists t, $|t|<\epsilon$, such that $K_t{}^n=1$ for a positive integer n. By Theorem 33, the corresponding surface S_t has a finite unramified covering surface $\tilde{S}_t$ with a trivial canonical bundle. Since $b_1(S_t)=1$, $\tilde{S}_t$ is non-Kähler and $b_1(\tilde{S}_t)\geqq 1$, Hence, by Theorem 19, $\tilde{S}_t$ is elliptic and consequently S_t is elliptic. Moreover, since $b_2(S_t)=0$, S_t belongs to the class VII_0. Therefore, by Theorem 37, S is either an elliptic surface or a Hopf surface. This contradicts the non-existence of curves on S.

ii) Now we consider the case in which $\dim H^1(\boldsymbol{C}(K))=1$. By (134), there exists on S a holomorphic vector field θ which does not vanish identically. A point z on S is called a singular point of θ if and only if θ vanishes at z. Since S contains no curve, θ has at most a finite number of singular points. The indices of the singular points are positive and their sum is equal to the Euler number c_2 of S, while c_2 is equal to zero. Consequently θ does not vanish at each point of S. Hence we can choose local complex coordinates (z_j,ζ_j) such that

$$\theta=\partial/\partial z_j.$$

We have

$$z_j=z_k+f_{jk}(\zeta_k), \tag{137}$$

where $f_{jk}(\zeta_k)$ is a holomorphic function of ζ_k. Since, by (126) and (133), $\dim H^0(d\mathcal{O}(K^{-1}))=\dim H^1(\boldsymbol{C}(K))=1$, there exists an element $\phi\in H^0(\ddot{d}\mathcal{O}(K^{-1}))$, $\phi\neq 0$, which has the form $\phi=\{dw_j\}$, where the w_j are local holomorphic functions and

$$dw_j=\kappa_{jk}dw_k, \qquad \kappa_{jk}=K^{-m_{jk}}. \tag{138}$$

Hence, applying θ to w_j and w_k, we get

$$\partial w_j/\partial z_j=\kappa_{jk}\partial w_k/\partial z_k.$$

It follows that $\partial w_j/\partial z_j=0$. Thus each function w_j depends only on ζ_j. It follows from (137) and (138) that

$$dz_j\wedge dw_j=\kappa_{jk}dz_k\wedge dw_k.$$

Since S contains no curve, this proves that the derivatives $dw_j/d\zeta_j$ never vanish. Consequently we can employ (z_j, w_j) as local coordinates. We denote by U_j the domain of (z_j, w_j). Obviously we may assume that U_j is a polycylinder: $|z_j| < \rho_j$, $|w_j| < \rho_j$, where $\rho_j > 0$, and that the intersections $U_j \cap U_k$ are connected. We infer from (137) and (138) that

$$\left\{\begin{array}{l} z_j + g_{jk}(w_j) = z_k, \\ w_j = \kappa_{jk} w_k + c_{jk}, \end{array}\right. \tag{139}$$

where $g_{jk}(w_j)$ is a holomorphic function of w_j and c_{jk} denotes a constant.

Let $g_{jk}''(w_j)$ denote the second derivative of $g_{jk}(w_j)$. We have

$$g_{ik}(w_i) = g_{ij}(w_i) + g_{jk}(w_j), \qquad \text{on } U_i \cap U_j \cap U_k.$$

Hence, using (138), we obtain

$$g_{ik}''(w_i) = g_{ij}''(w_i) + \kappa_{ij}^{-2} g_{jk}''(w_j).$$

This shows that the collection $\{g_{jk}''(w_j)\}$ forms a 1-cocycle with coefficients in $\mathcal{O}(K^2)$. Hence, by (125), $\{g_{jk}''(w_j)\}$ is cohomologous to zero, i. e., there exists a 0-cochain $\{h_j\}$ such that

$$g_{jk}''(w_j) = \kappa_{jk}^{-2} h_k - h_j, \qquad \text{on } U_j \cap U_k.$$

Applying θ to this, we get

$$\partial h_j / \partial z_j = \kappa_{jk}^{-2} \partial h_k / \partial z_k.$$

Hence, by (123), $\partial h_j / \partial z_j$ vanishes. Thus h_j is a holomorphic function $h_j(w_j)$ of w_j. We choose a holomorphic function $g_j(w_j)$ on each neighborhood U_j such that $g_j''(w_j) = h_j(w_j)$. Then we get

$$g_{jk}(w_j) = g_k(w_k) - g_j(w_j) + a_{jk} w_j + b_{jk}, \qquad \text{on } U_j \cap U_k,$$

where a_{jk} and b_{jk} are constants. Combining this with (139), we obtain

$$\left\{\begin{array}{l} z_j - g_j(w_j) + a_{jk} w_j + b_{jk} = z_k - g_k(w_k), \\ w_j = \kappa_{jk} w_k + c_{jk}. \end{array}\right. \tag{140}$$

We infer from (140) that

$$a_{ik} = a_{ij} + \kappa_{ij}^{-1} a_{jk}, \qquad \bar{c}_{ik} = \bar{c}_{ij} + \kappa_{ij}^{-1} \bar{c}_{jk},$$

provided that $U_i \cap U_j \cap U_k$ is non-empty. Thus $\{a_{jk}\}$ and $\{\bar{c}_{jk}\}$ are 1-cocycles with coefficients in $\boldsymbol{C}(K)$. $\{\bar{c}_{jk}\}$ is not cohomologous to zero. In fact, the cohomological relation: $c_{jk} = \kappa_{jk} c_k - c_j$ would imply that

$$w_j + c_j = \kappa_{jk}(w_k + c_k).$$

This contradicts (123). Since $\dim H^1(\boldsymbol{C}(K)) = 1$, $\{a_{jk}\}$ is cohomologous to $\{a \cdot \bar{c}_{jk}\}$, i. e.,

$$a_{jk} = a \cdot \bar{c}_{jk} + \bar{\kappa}_{jk} a_k - a_j, \tag{141}$$

where a is a constant. Letting

$$f_j = z_j - g_j(w_j) - a_j w_j,$$

we infer from (140) and (141) that

$$df_j + a\bar{w}_j dw_j = df_k + a\bar{w}_k dw_k.$$

Hence, setting

$$\sigma = df_j + d\bar{f}_j + ad\,|\,w_j\,|^2, \qquad \text{on each } U_j,$$

we obtain a d-closed real 1-form σ on S. σ is not d-cohomologous to zero. In fact, if a real-valued differentiable function τ on S satisfies $d\tau = \sigma$, then τ would not be able to attain its maximum. Let $\boldsymbol{R}$ denote the field of real numbers and let $T = \boldsymbol{R}/\boldsymbol{Z}$. Moreover let ω denote the period $\int_\beta \sigma$ of σ over a basic 1-cycle β on S. Then the formula

$$\mu\colon z \to \mu(z) = \omega^{-1} \int^z \sigma \,(\mathrm{mod}\, \boldsymbol{Z})$$

defines a differentiable map of S onto the circle T, since $b_1(S)$ is equal to 1. We infer readily that the map μ determines on S a structure of differentiable fibre bundle over T whose fibre is a compact connected orientable differentiable manifold X of topological dimension 3. Therefore S is obtained from a cartesian product $[0,1] \times X$ by the identification: $(1, x) = (0, h(x))$, where $h\colon x \to h(x)$ is a differentiable automorphism of X. It is obvious that $H^1(S, \boldsymbol{Z})$ is induced from $H^1(T, \boldsymbol{Z})$ by the map μ. Hence the complex line bundle F is obtained from $[0,1] \times X \times \boldsymbol{C}$ by the identification: $(1, x, F\zeta) = (0, h(x), \zeta)$. Using these descriptions of S and F we infer readily the following result: Let h^* denote the linear transformation of $H^1(X, \boldsymbol{C})$ induced by the differentiable automorphism h. Then the dimension of $H^1(S, \boldsymbol{C}(F))$ is equal to the number of linearly independent solutions of the linear equation:

$$h^*\xi = F\xi, \qquad \xi \in H^1(X, \boldsymbol{C}),$$

provided that $F \neq 1$. Since h^* transforms $H^1(X, \boldsymbol{Z})$ onto itself, h^* is represented by an integral unimodular matrix U. Let

$$\chi(\lambda) = \det(\lambda - U).$$

The above result implies that $\dim H^1(S, \boldsymbol{C}(F))$ is positive if and only if $\chi(F)$ vanishes, provided that $F \neq 1$. On the other hand, since $\dim H^1(\boldsymbol{C}(K)) = 1$, it follows from (127), (129), (133) and (134) that

$$\dim H^1(S, \boldsymbol{C}(F)) = \begin{cases} 1, & \text{for } F = 1, K, K^{-1}, \\ 0 & \text{otherwise.} \end{cases}$$

Hence the polynomial $\chi(\lambda)$ has the form:

$$\chi(\lambda) = (\lambda - K)^m (\lambda - \bar{K})^m (\lambda - 1)^n, \qquad m \geqq 1, n \geqq 0.$$

This proves that $K + \bar{K}$ is a rational integer, while $\bar{K}K = 1$. Consequently K^{12} is equal to 1, q. e. d.

Consider a surface S with a volume-preserving complex structure. By the above theorem, S has a finite unramified normal covering surface $\tilde{S}$ with a trivial canonical bundle. By Theorem 19, $\tilde{S}$ is either a K3 surface or a surface of the form $\boldsymbol{C}^2/G$, where G is either a discrete subgroup of the additive group $\boldsymbol{C}^2$ (then $\tilde{S}$ is a complex torus) or a properly discontinuous non-abelian group of affine transformations of $\boldsymbol{C}^2$.

First we examine the case in which $\tilde{S}$ is a K3 surface. Let m denote the number of sheets of the covering surface $\tilde{S}$ of S. Then $mc_2(S)$ is equal to $c_2(\tilde{S}) = 24$, while, since $c_1{}^2(S)$ vanishes $c_2(S)$ is divisible by 12. Hence m is equal to either 1 or 2. If $m = 1$, then S is obviously a K3 surface. By an *Enriques surface* we shall mean a surface with numerical characters $p_g = q = 0$, $P_2 = 1$ (see Enriques [1], Cap. VII, § 1). We infer readily that, if $m = 2$, then S is an Enriques surface. By Theorem 13, every K3 surface is simply connected. Hence, in this case, the K3 surface $\tilde{S}$ is the universal covering manifold of S.

Now we consider the case in which $\tilde{S} = \boldsymbol{C}^2/G$. Obviously S is represented as a quotient surface: $S = \boldsymbol{C}^2/\mathfrak{G}$, where $\mathfrak{G}$ is a properly discontinuous group of complex analytic automorphisms of $\boldsymbol{C}^2$ which contains G as a normal subgroup of finite index. We shall prove that $\mathfrak{G}$ *consists of affine transformations of* $\boldsymbol{C}^2$. Take an element $g \in G$ and write it in the form

$$g: z = (z_1, z_2) \to g(z) = (g_1(z), g_2(z)).$$

g induces a complex analytic automorphism of $\tilde{S} = \boldsymbol{C}^2/G$ which, in turn, induces a linear transformation of the space $H^0(\tilde{S}, \Omega^r)$ of holomorphic r-forms on $\tilde{S}$.

In the case in which $\tilde{S}$ is a complex torus, $H^0(\tilde{S}, \Omega^1)$ is spanned by dz_1 and dz_2. Hence we have

$$dg_\nu(z) = \alpha_{\nu 1} dz_1 + \alpha_{\nu 2} dz_2, \qquad \nu = 1, 2,$$

where $\alpha_{\nu 1}$ and $\alpha_{\nu 2}$ are constants. It follows that

$$g_\nu(z) = \alpha_{\nu 1} z_1 + \alpha_{\nu 2} z_2 + \text{constant}, \qquad \nu = 1, 2.$$

In the case in which G is non-abelian, G is generated by four affine transformations

$$A_j\colon (z_1, z_2) \to (z_1 + \alpha_j, z_2 + \bar{\alpha}_j z_1 + \beta_j), \qquad j = 1, 2, 3, 4,$$

where $\alpha_1 = \alpha_2 = 0$ and $\bar{\alpha}_3 \alpha_4 - \bar{\alpha}_4 \alpha_3 \neq 0$. $H^0(\tilde{S}, \Omega^1)$ and $H^0(\tilde{S}, \Omega^2)$ are spanned by dz_1 and by $dz_1 \wedge dz_2$, respectively (see § 6). Hence we have

$$dg_1(z) = \sigma dz_1, \qquad dg_1(z) \wedge dg_2(z) = \kappa dz_1 \wedge dz_2,$$

where σ and κ are constants. It follows that

$$g_1(z) = \sigma z_1 + \gamma, \qquad g_2(z) = \sigma^{-1} \kappa z_2 + h(z_1),$$

where γ is a constant and $h(z_1)$ is a holomorphic function of z_1. Since the index of G in $\mathfrak{G}$ is finite, there exists a positive integer n such that $g^n \in G$. It follows that σ and κ are roots of unity. Obviously $A_j^* = g A_j g^{-1}$ belongs to G and has the form

$$A_j^*\colon (z_1, z_2) \to (z_1 + \alpha_j^*, z_2 + \bar{\alpha}_j^* z_1 + \beta_j^*).$$

We infer from the relation $A_j^* g = g A_j$ that

$$h'(z_1 + \alpha_j) - h'(z_1) = (1 - \sigma^{-1}\kappa) \bar{\alpha}_j,$$

where $h' = dh/dz_1$. This implies that $h''(z_1)$ is invariant under the translations: $z_1 \to z_1 + \alpha_j$. Hence $h'' = h''(z_1)$ is a constant and

$$h'' \cdot \alpha_j = (1 - \sigma^{-1}\kappa) \bar{\alpha}_j, \qquad j = 3, 4.$$

This proves that $h'' = 0$, $1 - \sigma^{-1}\kappa = 0$, and consequently,

$$g_1(z) = \sigma z_1 + \gamma, \qquad g_2(z) = z_2 + h_1 z_1 + h_2,$$

where h_1 and h_2 are constants. Thus we obtain the following

THEOREM 39. *If a surface S has a volume-preserving complex structure, then the universal covering manifold $\mathfrak{U}$ of S is either a* K3 *surface or $\boldsymbol{C}^2$. If $\mathfrak{U}$ is a* K3 *surface, then S is either a* K3 *surface or an Enriques surface. If $\mathfrak{U}$ is $\boldsymbol{C}^2$, then S is a surface of the form $\boldsymbol{C}^2/G$, where G is a properly dis-*

continuous group of volume-preserving affine transformations without fixed points of $\boldsymbol{C}^2$.

12. Discontinuous groups of volume-preserving complex analytic automorphisms of $\boldsymbol{C}^2$. Let $\mathfrak{G}$ be a properly discontinuous group of complex analytic automorphisms of $\boldsymbol{C}^2$ without fixed points. Assume that $\mathfrak{G}$ has a bounded fundamental domain. Then $\boldsymbol{C}^2/\mathfrak{G}$ is a surface with volume-preserving complex structure. Hence, by Theorem 39, $\boldsymbol{C}^2/\mathfrak{G}$ can be represented in the form $\boldsymbol{C}^2/G$, where G is properly discontinuous group of volume-preserving affine transformations of $\boldsymbol{C}^2$. This proves the existence of a volume-preserving complex analytic automorphism h of $\boldsymbol{C}^2$ such that $h\mathfrak{G}h^{-1} = G$. Thus we obtain the following theorem:

THEOREM 40. *If a properly discontinuous group $\mathfrak{G}$ of volume-preserving complex analytic automorphisms of $\boldsymbol{C}^2$ is free from fixed points and has a bounded fundamental domain, then there exists a volume-preserving complex analytic automorphisms h of $\boldsymbol{C}^2$ such that $h\mathfrak{G}h^{-1}$ becomes a properly discontinuous group of volume-preserving affine transformations of $\boldsymbol{C}^2$.*

REFERENCES.

[1] F. Enriques, *Le Superficie Algebriche*, Bologna, 1949.
[2] K. Kodaira, "On compact complex analytic surfaces, I," *Annals of Mathematics*, vol. 71 (1960), pp. 111-152.
[3] ———, "On compact analytic surfaces, II-III," *Annals of Mathematics*, vol. 77 (1963), pp. 563-626, vol. 78 (1963), pp. 1-40.
[4] ———, "On stability of compact submanifolds of complex manifolds," *American Journal of Mathematics*, vol. 85 (1963), pp. 79-94.
[5] ———, "On the structure of compact complex analytic surfaces, I," *American Journal of Mathematics*, vol. 86 (1964), pp. 751-798.
[6] ———, L. Nirenberg and D. C. Spencer, "On the existence of deformations of complex analytic structures," *Annals of Mathematics*, vol. 68 (1958), pp. 450-459.
[7] ——— and D. C. Spencer, "On deformations of complex structures, I-II," *Annals of Mathematics*, vol. 67 (1958), pp. 328-466.
[8] ——— and ———, "A theorem of completeness for complex analytic fibre spaces," *Acta Mathematica*, vol. 100 (1958), pp. 281-294.
[9] M. S. Lattès, "Sur les formes réduites des transformations ponctuelles à deux variables," *Comptes Rendus*, vol. 152 (1911), pp. 1566-1569.
[10] S. Sternberg, "Local contractions and a theorem of Poincaré," *American Journal of Mathematics*, vol. 79 (1957), pp. 809-824.

THE HAUSDORFF DIMENSION OF SINGULAR SETS OF PROPERLY DISCONTINUOUS GROUPS.

By A. F. Beardon.

---

1. Introduction. Let G be a properly discontinuous group of transformations, each element of which maps euclidean N-dimensional space onto itself. Associated with G is a set E, called the singular set of G, on which G is not properly discontinuous.

The object of this paper is to extend certain well known results which relate the density of E to the generators of G. The main result is that the singular set of a Kleinian group has positive Hausdorff dimension.

2. Hausdorff measures. Let E be a compact subset of euclidean N-dimensional space, E^N. Define

$$m_{\alpha,\delta}(E) = \inf \sum_i |I_i|^\alpha$$

where the infimum is taken over all coverings of E by sequences $\{I_i\}$ of sets I_i with diameter $|I_i|$ less than δ. Further, define

$$m_\alpha(E) = \sup_{\delta>0} m_{\alpha,\delta}(E)$$

and this is the Hausdorff α-dimensional measure of E ([4]).

The Hausdorff dimension of E is defined to be the unique non-negative number $d(E)$ such that

$$m_\alpha(E) = 0 \text{ if } \alpha > d(E)$$

and

$$m_\alpha(E) = \infty \text{ if } 0 \leqq \alpha < d(E).$$

Another measure of the density of E is the logarithmic capacity (e.g. [8]) which we shall denote by $C(E)$. The following result is well known.

Theorem 1 ([8]). *Let E be a compact subset of E^2. Then $d(E) > 0$ implies $C(E) > 0$.*

3. Spherical Cantor sets. We shall use the notion of a spherical Cantor set ([10]), the definition being given here for completeness.

Received April 19, 1965.

Definition 1. A subset E of E^N is a spherical Cantor set if and only if E can be expressed in the form

$$E = \bigcap_{n=1}^{\infty} \bigcup_{i_1,\cdots,i_n=1}^{K} I(i_1,\cdots,i_n)$$

where K is a positive integer greater than one and the $I(i_1,\cdots,i_n)$ are closed N-dimensional spheres of radius $r(i_1,\cdots,i_n)$ satisfying

(a) $I(i_1,\cdots,i_n) \supset I(i_1,\cdots,i_n,i_{n+1}) \qquad (i_{n+1}=1,\cdots,K)$

(b) $I(1),\cdots,I(K)$ are mutually disjoint,

(c) there exists a constant A, $1 > A > 0$, such that

$$r(i_1,\cdots,i_n,i_{n+1}) \geqq A r(i_1,\cdots,i_n) \qquad (i_{n+1}=1,\cdots,K)$$

and

(d) there exists a constant B, $1 > B > 0$, such that

$$\rho[I(i_1,\cdots,i_n,j), I(i_1,\cdots,i_n,k)] \geqq B r(i_1,\cdots,i_n)$$

where

$$\rho[S,T] = \inf\{|s-t|; s \in S, t \in T\}$$

and $j,k = 1,\cdots,K$; $j \neq k$.

We quote the following results on spherical Cantor sets.

THEOREM 2 ([2]). *Let E be a spherical Cantor set such that there exist constants A_j $(j=1,\cdots,K;\ A_j \geqq 0)$ satisfying*

$$r(i_1,\cdots,i_n,j) \geqq A_j r(i_1,\cdots,i_n).$$

Then $d(E) = d$ satisfies

$$A_1{}^d + \cdots + A_K{}^d \leqq 1.$$

COROLLARY. *Since for each j, $A_j \geqq A$ it follows that*

$$d(E) \geqq \frac{-\log K}{\log A} > 0$$

and so every spherical Cantor set has positive Hausdorff dimension.

THEOREM 3 ([2]). *There exists a function $F(K,N)$ defined for $K,N = 2,3,\cdots$ such that if E is a spherical Cantor set in E^N with K as in Definition 1 then*

$$d(E) \leqq F(K,N) < N.$$

In particular, $d(E) < N$.

THEOREM 4 ([2]). *Let E be a spherical Cantor set and C a constant satisfying*

(a) $$1 > C > 0$$

and

(b) $$r(i_1, \cdots, i_n) = 0(C^n) \text{ as } n \to \infty.$$

Then

$$d(E) \leqq \frac{-\log K}{\log C}.$$

We note here that Theorem 4 is applicable if for some C, $1 > C > 0$,

$$r(i_1, \cdots, i_n, i_{n+1}) \leqq Cr(i_1, \cdots, i_n).$$

4. Inversion groups. The general properties of properly discontinuous groups are discussed in [3] and [5]. In this section we restrict the discussion of such groups to the following special case.

For $j = 1, \cdots, K+1$ let S_j be an N-dimensional sphere in E^N with centre $a(j)$ and radius $r(j)$. Suppose that $S_1, \cdots, S_{K+1}$ are mutually disjoint and for any $x \in E^N$ denote by $I_j(x)$ the inverse joint of x with respect to S_j. It is easily shown that $I_1, \cdots, I_{K+1}$ generate a properly discontinuous group with the complement of $\bigcup_j S_j$ as a fundamental region. Such a group is called an inversion group and is finitely generated. If G is an inversion group with $K = 1$ then the singular set of G consists of two points only.

Define $\sum(n)$ to be the set of all sequences $i_1, \cdots, i_n$ where

$$i_1, \cdots, i_n \in \{1, \cdots, K+1\} \text{ and } i_1 \neq i_2, \cdots, i_{n-1} \neq i_n.$$

Further, define

$$I(j) = S_j$$

for $j = 1, \cdots, K+1$ and

$$I(i_1, \cdots, i_n) = I_{i_1} \cdots I_{i_{n-1}}(S_{i_n})$$

where $i_1, \cdots, i_n \in \sum(n)$, $n \geqq 2$ and $I_s(I_t(x)) = I_s I_t(x)$.

LEMMA 1. *Let E be the singular set of the inversion group G. Then in the above notation*

$$E = \bigcap_{n=1}^{\infty} \bigcup_{i_1, \cdots, i_n \in \Sigma(n)} I(i_1, \cdots, i_n),$$

Proof. From the geometry of inversions together with the fact that S_j are mutually external spheres it is clear that the compact sets

$$H_n = \cup\, I(i_1,\cdots,i_n) \qquad\qquad (i_1,\cdots,i_n \in \Sigma(n))$$

satisfy

$$H_n \supset H_{n+1} \qquad\qquad n = 1, 2, \cdots$$

and from the definition of G it is immediate that for each integer n, G is properly discontinuous in the complement of H_n. Thus

$$E \subset \bigcap_{n=1}^{\infty} H_n.$$

Since

$$I_{i_1}\cdots I_{i_n}(\infty) \in I_{i_1}\cdots I_{i_{n-1}}(S_{i_n})$$

it follows that for every $I(i_1,\cdots,i_n)$ there exists a $g \in G$ with

$$g(\infty) \in I(i_1,\cdots,i_n). \tag{1}$$

If $a(i_1,\cdots,i_n)$, $r(i_1,\cdots,i_n)$ denote the centre and radius respectively of $I(i_1,\cdots,i_n)$ the definition of inversion in S_{i_1} yields

$$r(i_1,\cdots,i_n) = \frac{r(i_1)^2 r(i_2,\cdots,i_n)}{|a(i_1) - a(i_2,\cdots,i_n)|^2 - r(i_2,\cdots,i_n)^2} \tag{2}$$

and hence

$$\begin{aligned} r(i_1,\cdots,i_n) &\leqq \frac{r(i_1)^2 r(i_2,\cdots,i_n)}{[|a(i_1) - a(i_2)| - r(i_2)]^2} \\ &\leqq \lambda^2 r(i_2,\cdots,i_n) \end{aligned}$$

where

$$\begin{aligned} \lambda &= \max_{s \neq t} \frac{r(s)}{|a(s) - a(t)| - r(t)} \\ &< 1. \end{aligned}$$

This implies that

$$r(i_1,\cdots,i_n) \leqq \lambda^{2n-2} r(i_n) \tag{3}$$

and hence that

$$r(i_1,\cdots,i_n) \rightarrow 0 \text{ as } n \rightarrow \infty \tag{4}$$

uniformly over all sequences $\{i_n\}$.

From (1) and (3) we deduce that any point x of the form

$$\{x\} = \bigcap_{n=1}^{\infty} I(i_1, \cdots, i_n)$$

(where $\{i_n\}$ is some infinite sequence with $i_1, \cdots, i_n \in \sum(n)$) is a limit point of the set $\{g(\infty) : g \in G\}$ and hence $x \in E$.

Thus

$$E = \bigcap_{n=1}^{\infty} H_n$$

as required.

THEOREM 5. *Let E be the singular set of an inversion group. Then*

$$d(E) \leqq \frac{-\log K}{2 \log \lambda}$$

where λ is as defined above.

Proof. It follows from (3) that

$$r(i_1, \cdots, i_n) = 0(\lambda^{2n}) \text{ as } n \to \infty$$

and an application of Theorem 4 yields the desired result. We next prove the basic result of this section.

THEOREM 6. *Let G be an inversion group with singular set E and suppose that $K \geqq 2$. Then $E_j = E \cap S_j$ is a spherical Cantor set, $(j = 1, \cdots, K + 1)$.*

Proof. By Lemma 1 it follows that

$$E_j = \bigcap_{n=1}^{\infty} \bigcup_{i_1, \cdots, i_n \in \Sigma(n)} I(i_1, \cdots, i_n)$$

where $i_1 = j$. Apart from relabelling this is the form required of E_j by Definition 1; it is more convenient, however, to retain this notation. The conditions of Definition 1 are trivially satisfied except (c) and (d) and (c) is established first.

By elementary geometry we can estimate

$$r(i_1, \cdots, i_n, i_{n+1})\, r(i_1, \cdots, i_n)^{-1}$$

in terms of

$$r(i_2, \cdots, i_n, i_{n+1})\, r(i_2, \cdots, i_n)^{-1}$$

and hence ultimately in terms of

$$r(i_n, i_{n+1})\, r(i_n)^{-1}$$

which can easily be computed.

Write

$$\theta(i_1,\cdots,i_n,i_{n+1})=r(i_1,\cdots,i_n,i_{n+1})r(i_1,\cdots,i_n)^{-1}$$

and we can immediately deduce from (2) of Lemma 1 that

$$\frac{\theta(i_1,\cdots,i_n,i_{n+1})}{\theta(i_2,\cdots,i_n,i_{n+1})}=\frac{|a(i_1)-a(i_2,\cdots,i_n)|^2-r(i_2,\cdots,i_n)^2}{|a(i_1)-a(i_2,\cdots,i_n,i_{n+1})|^2-r(i_2,\cdots,i_n,i_{n+1})^2}.$$

From geometrical considerations we have

$$\begin{aligned}[|a(i_1)-a(i_2,\cdots,i_n)|-r(i_2,\cdots,i_n)]^2&\\ \leqq|a(i_1)-a(i_2,\cdots,i_{n+1})|^2&-r(i_2,\cdots,i_{n+1})^2\\ \leqq[|a(i_1)-a(i_2,\cdots,i_n)|&+r(i_2,\cdots,i_n)]^2\end{aligned}$$

and so

$$\mu^{-1}\theta(i_2,\cdots,i_{n+1})\leqq\theta(i_1,\cdots,i_{n+1})\leqq\mu\theta(i_2,\cdots,i_{n+1}) \tag{5}$$

where

$$\begin{aligned}\mu&=\frac{|a(i_1)-a(i_2,\cdots,i_n)|+r(i_2,\cdots,i_n)}{|a(i_1)-a(i_2,\cdots,i_n)|-r(i_2,\cdots,i_n)}\\ &=1+\frac{2r(i_2,\cdots,i_n)}{|a(i_1)-a(i_2,\cdots,i_n)|-r(i_2,\cdots,i_n)}.\end{aligned}$$

With the notation

$$d_{st}=|a(s)-a(t)| \qquad (s\neq t)$$

and

$$d=\min\{d_{st}-r(s):s\neq t\}$$

we can write

$$\begin{aligned}\mu&\leqq1+2d^{-1}r(i_2,\cdots,i_n)\\ &\leqq1+2d^{-1}\lambda^{2n-4}r(i_n)\end{aligned}$$

where λ is as defined in Lemma 1. Writing

$$Q=\frac{2}{d\lambda^4}\max\{r(s)\}$$

we have

$$\mu\leqq1+Q\lambda^{2n}.$$

Repeated applications of (5) together with this estimate of μ yield

$$P^{-1}\theta(i_n,i_{n+1})\leqq\theta(i_1,\cdots,i_n,i_{n+1})\leqq P\theta(i_n,i_{n+1}) \tag{6}$$

where

$$P = \prod_{n=2}^{\infty} (1 + Q\lambda^{2n})$$

and this converges since $\lambda < 1$. Since

$$\{\theta(i_n, i_{n+1}) : i_n, i_{n+1} \in \sum(2)\}$$

is a finite set we have shown that $\theta(i_1, \cdots, i_{n+1})$ is bounded below by a positive constant and this is (c) of Definition 1.

Using a similar argument it is possible to show that

$$\frac{\rho[I(i_1, \cdots, i_n, s), I(i_1, \cdots, i_n, t)]}{r(i_1, \cdots, i_n)} \geqq \mu^{-2} \frac{\rho[I(i_2, \cdots, i_n, s), I(i_2, \cdots, i_n, t)]}{r(i_2, \cdots, i_n)}$$

and repeated applications of this inequality establish (d) of Definition 1. This completes the proof of Theorem 6.

THEOREM 7. *Let G be an inversion group with $K \geqq 2$ and with singular set E. Then*

(a) $$0 < \frac{\log K}{\log (P/\theta')} \leqq d(E)$$

and

(b) $$P\theta'' < 1 \quad \textit{implies} \quad d(E) \leqq \frac{-\log K}{\log (P\theta'')}$$

where

$$\theta' = \min\{\theta(s,t)\}$$
$$\theta'' = \max\{\theta(s,t)\} \qquad (s \neq t)$$

and where P is as defined in the proof of Theorem 6. In particular, E has positive Hausdorff dimension.

Proof. From (6) we see that for all $i_1, \cdots, i_{n+1} \in \sum(n+1)$,

$$P^{-1}\theta' \leqq \theta(i_1, \cdots, i_{n+1}) \leqq P\theta''.$$

The lower bound follows from the corollary of Theorem 2 and Theorem 6 and the upper bound follows from Theorems 4 and 6.

If the spheres S_j are such that

$$r(1) = \cdots = r(K+1) = 1$$

the bounds of Theorem 6 can be considerably simplified. In this case define

$$e = \min\{d_{st} : s \neq t\}, \qquad f = \max\{d_{st} : s \neq t\}$$

and estimate P as follows:

$$
\begin{aligned}
P &= \prod_{n=0}^{\infty}\left(1+\frac{2}{(e-1)^{2n+1}}\right) \\
&\leqq \exp\left\{\sum_{n=0}^{\infty}\frac{2}{(e-1)^{2n+1}}\right\} \\
&= \exp\left\{\frac{2(e-1)}{e(e-2)}\right\} \\
&< \exp\left\{\frac{2}{e-2}\right\}.
\end{aligned} \tag{7}
$$

Thus in this case

$$\frac{\theta'}{\exp\left(\frac{2}{e-2}\right)} \leqq \theta(i_1,\cdots,i_n,i_{n+1}) \leqq \theta'' \exp\left(\frac{2}{e-2}\right).$$

Since

$$\theta(s,t)=\frac{1}{d_{st}{}^2-1}$$

it follows that

$$
\begin{aligned}
(e^2-1)^{-1} &\geqq \theta(s,t) \\
&\geqq d_{st}{}^{-2} \\
&\geqq f^{-2}
\end{aligned} \tag{8}
$$

and hence that

$$\left[f^2\exp\left(\frac{2}{e-2}\right)\right]^{-1} \leqq \theta(i_1,\cdots,i_n,i_{n+1}) \leqq (e^2-1)^{-1}\exp\left(\frac{2}{e-2}\right).$$

These estimates lead directly to a proof of the following corollary of Theorem 7.

COROLLARY. *Let G be an inversion group with $K \geqq 2$, singular set E and $r(1)=\cdots=r(K+1)=1$. Then*

$$d(E) \geqq \frac{\log K}{2\log f + \left(\frac{2}{e-2}\right)} > 0 \tag{a}$$

and

$$d(E) \leqq \frac{\log K}{\log(e^2-1) - \left(\frac{2}{e-2}\right)} \tag{b}$$

provided that

$$(e^2-1) > \exp\left(\frac{2}{e-2}\right).$$

THEOREM 8. *Let G be an inversion group in E^N, $(N \geqq 2)$, with K defined as in Definition 1 and with singular set E. Then*

$$d(E) \leqq F(K,N) < N$$

where $F(K,N)$ is a function of the positive integers N and K only. In particular, $d(E) < N$.

The proof follows directly from Theorems 3 and 6.

It is of interest to study the range of values of $d(E)$. From Theorem 7 we see that $d(E) > 0$ and Theorem 5 implies that this lower bound is the best possible. From Theorem 8 we know that $d(E) < N$ but in this case the best possible bound is not known. However, we can prove the following result.

THEOREM 9. *There exists an inversion group with residual set E in E^N such that $d(E) > N/2$.*

Proof. We shall construct a sequence of inversion groups G_n with singular sets E_n such that for $n = 1, 2, \cdots$

$$E_n \subset E_{n+1}. \tag{9}$$

The sequence $\{d(E_n)\}$ increases monotonically to a finite limit, say d, and Theorem 8 will be proved if we can show that $d > N/2$.

Let L_n denote the set

$$\{(x_1, \cdots, x_N) : x_j = 0, 3, 6, \cdots, 6n; j = 1, \cdots, N\}$$

and construct spheres of unit radius about each point of L_n. Call this system of spheres Q_n and let G_n be the inversion group associated with these spheres. Since G_n is a subgroup of G_{n+1} the relation (9) above is established.

Now let K_p be the value of K corresponding to G_p and consider the following discussion as applied to G_p.

Since $e = 3$ and $r(s) = 1$ $(s = 1, \cdots, K_p + 1)$ it follows from (6), (7) and (8) that

$$\theta(i_1, \cdots, i_n, i_{n+1}) \geqq (8 d_{i_n i_{n+1}}^2)^{-1}. \tag{10}$$

With a view to using Theorem 2 we shall estimate the numbers $d_{i_n i_{n+1}}$ for a fixed sequence $i_1, \cdots, i_n \in \sum(n)$ and for $i_{n+1} = 1, \cdots, K_p + 1, \neq i_n$. By virtue of the definition of L_p the sequence

$$\{d^2_{i_n i_{n+1}} : i_{n+1} = 1, \cdots, K_p + 1; \neq i_n\}$$

always contains as a subsequence (possibly reordered) the sequence

$$\{t_1^2 + \cdots + t_N^2 : t_j = 3, 6, \cdots, 3p; j = 1, \cdots, N\}.$$

Thus the sequence of numbers

$$\{(8d^2_{i_n i_{n+1}})^{-1} : i_{n+1} = 1, \cdots, K_p + 1; \neq i_n\}$$

always contains the sequence

$$\{\frac{1}{8(t_1^2 + \cdots + t_N^2)} : t_j = 3, 6, \cdots, 3p; j = 1, \cdots, N\}$$

and so by (10) we may take p^N of the constants A_j of Theorem 2 to be the p^N numbers

$$\frac{1}{8(t_1^2 + \cdots + t_N^2)} \qquad t_j \in \{3, 6, \cdots, 3p\}$$

and the remaining A_j's to be zero. With this choice Theorem 2 implies that

$$1 \geqq \sum_{t_1,\cdots,t_N=1}^{p} \left\{\frac{1}{72(t_1^2 + \cdots + t_N^2)}\right\}^{d(E_p)}$$

and since $d(E_p) \leqq d$ that

$$\sum_{t_1,\cdots,t_N=1}^{p} \left\{\frac{1}{72(t_1^2 + \cdots + t_N^2)}\right\}^{d} \leqq 1.$$

It follows that

$$\sum_{t_1,\cdots,t_N=1}^{\infty} \frac{1}{(t_1^2 + \cdots + t_N^2)^d}$$

is convergent and the proof is completed by showing that this implies $2d > N$.

For each $m = 1, 2, \cdots$ we have

$$\sum_{t_1,\cdots,t_N=2^m+1}^{2^{m+1}} (t_1^2 + \cdots + t_N^2)^{-d} \geqq 2^{mN}(4N2^{2m})^{-d}$$

and so

$$\sum_{m=1}^{\infty} 2^{mN}(4N2^{2m})^{-d}$$

converges which clearly implies $2d > N$.

5. Properly discontinuous groups in E^2.

In this section we restrict our attention to transformations of the extended complex plane into itself. The definition and details of properly discontinuous groups of linear transformations in the complex plane can be found in [3] and [5].

Let $C_1, \cdots, C_{2p}$ be $2p$ disjoint and mutually external circles and let T_j be a linear transformation mapping the exterior of C_j onto the interior of C_{j+p} $(j = 1, \cdots, p)$. The transformations $T_1, \cdots, T_p$ generate a properly discontinuous group G, known as a Schottky group, and the infinite domain bounded by the circles $C_1, \cdots, C_{2p}$ can be taken as a fundamental region for G.

In the case when the circles C_j, C_{j+p} are the isometric circles of T_j, T_j^{-1} respectively (for the definition and details of isometric circles see [3], p. 25) we follow Akaza ([1]) and call G a Schottky group of the Ford type. Since the radii of the isometric circles of T_j, T_j^{-1} are necessarily equal it follows that in a Schottky group of the Ford type the radii of C_j, C_{j+p} are the same.

We shall discuss here results relating to the Hausdorff dimension of the singular set of a properly discontinuous group.

THEOREM 10 ([6]). *Let E be the singular set of a non-cyclic Schottky group. Then $C(E) > 0$.*

Myrberg extended this result by showing that if G is a properly discontinuous group whose singular set has at least three points then G contains a non-cyclic Schottky group as a subgroup. Thus Myrberg proved the following result.

THEOREM 11 ([6]). *Let E be the singular set of a properly discontinuous group. If E has more than two points, then $C(E) > 0$.*

We shall use the following theorem to derive an extension of Theorem 11.

THEOREM 12 ([9], p. 109). *Let E be the singular set of a finitely generated non-cyclic Schottky group. Then that part of E contained in C_j $(j = 1, \cdots, 2p)$ is a spherical Cantor set.*

Theorem 1, the corollary of Theorem 2, Theorem 12 and Myrberg's own argument yield a proof of the following extension of Theorem 11.

THEOREM 13. *Let E be the singular set of a properly discontinuous group. If E has more than two points then $d(E) > 0$.*

THEOREM 14 ([7]). *Let E be the singular set of a finitely generated Schottky group. Then $m_2(E) = 0$.*

Theorems 3 and 12 yield a proof of the following extension of Theorem 14.

THEOREM 15. *There exists a function $F(p)$ defined for all positive* *integers p such that if E is the singular set of a Schottky group with p generators then*

$$d(E) \leqq F(p) < 2.$$

In particular, $d(E) < 2$.

6. Groups of the Ford type. It is well known (e.g. [3] p. 28) that a generating transformation T_j of a Schottky group of the Ford type is equivalent to the transformation obtained by inverting in C_j, reflecting into C_{j+p} and finally rotating about the centre of C_{j+p}. The rotation will be absent if T_j is a non-loxodromic transformation but in any case it is only the inversion which alters the distance between two points.

The singular set of a Schottky group is constructed in an analogous manner to the set E described in Lemma 1 (bilinear mappings replacing inversions). It is easily checked that the relation between inversions and bilinear mappings mentioned above implies that all of the estimates obtained earlier for the transformation $I_1(x), \cdots, I_{K+1}(x)$ are, in fact, valid for the bilinear mappings $T_1, T_1^{-1}, \cdots, T_p, T_p^{-1}$.

It follows that if G is a Schottky group of the Ford type and if each of the radii of $C_1, \cdots, C_{2p}$ is one we can deduce that the singular set E of G has dimension $d(E)$ satisfying

$$d(E) \geqq \frac{\log(2p-1)}{2\log f + 2(e-2)^{-1}} > 0$$

where e and f are the minimum and maximum distances respectively between the centres of C_s and C_t $(s \neq t;\, s, t = 1, \cdots, 2p)$. It is noteworthy that this bound is independent of the exact form of the generators $T_1, \cdots, T_p$ of G.

THEOREM 16 ([1]). *There exists a finitely generated Schottky group with singular set E which is such that* $d(E) > 1$.

Using the above relation between the inversion in C_j and the generators $T_1, \cdots, T_p$ of G it can be seen that Theorem 16 can be proved in a similar manner to Theorem 9.

7. Fuchsian groups. A Fuchsian group G is defined to be a properly discontinuous group of bilinear transformations each element of which leaves the real line invariant.

It is known (e.g. [3], p. 68) that the singular set E of G is either the whole real line (G is then a Fuchsian group of the first kind) or a nowhere dense subset of the real line (G is then a Fuchsian group of the second kind).

For any positive integer n the transformation

$$T_n(z) = \frac{4n(n+1) - (2n+1)z}{z - (2n+1)}$$

leaves the real axis invariant and has isometric circle

$$|z - (2n+1)| = 1.$$

Further, the isometric circle of $T_n^{-1}(z)$ is

$$|z + (2n+1)| = 1.$$

The group G_p generated by $T_2(z), T_4(z), \cdots, T_{2p}(z)$ can be regarded both as a Schottky group and as a Fuchsian group of the second kind. The discussion as applied earlier to a general Schottky group is applicable here; thus we have the following result.

THEOREM 17. *There exists a finitely generated Fuchsian group of the second kind with a singular set E such that $d(E) > \frac{1}{2}$.*

Appendix. (Added January 20, 1966) Let E be any bounded set in the complex plane. The local Hausdorff dimension of E at z is defined to be

$$d_z(E) = \lim_{\delta \to 0} d(E \cap S(z,\delta)) = \inf_{\delta > 0} d(E \cap S(z,\delta)),$$

where $S(z,\delta)$ is the closed disc with centre z and radius δ.

In order to measure sets containing $z = \infty$ we define the quantities $M_\alpha(E)$, $D(E)$ and $D_z(E)$ as for $m_\alpha(E)$, $d(E)$ and $d_z(E)$ respectively except that the chordal metric replaces the euclidean metric throughout these definitions. As in the original definition, if $\alpha < \beta$ and $M_\alpha(E) < +\infty$ then $M_\beta(E) = 0$ and hence

$$D(E) = \inf\{\alpha \colon M_\alpha(E) = 0\}.$$

LEMMA 2. (i) *Let A be any set which is bounded in the euclidean metric. Then $d(A) = D(A)$ and for any finite a, $d_a(A) = D_a(A)$.* (ii) *Let A be any set and let T be any non-singular bilinear transformation. Then for any a in the extended plane, $d_a(A) = D_{T(a)}(TA)$.*

Proof. Part (i) is easily proved using standard techniques. To prove (ii) we note that since T is a bilinear mapping,

$$\delta[T(z), T(z')] \leqq K\delta[z, z']$$

where $\delta[z, z']$ is the chordal metric and K depends only on the coefficients of T. This inequality together with the fact that if $\{I_i\}$ is a covering of A then $\{T(I_i)\}$ is a covering of $T(A)$ yields a simple proof of (ii).

THEOREM 18. *Let G be any properly discontinuous group of bilinear transformations with singular set E. Then*

$$D_z(E) = \begin{cases} D(E) & \text{if } z \in E \\ 0 & \text{if } z \notin E. \end{cases}$$

Proof. Since E is closed $D_z(E) = 0$ if $z \notin E$. The result is obvious if E contains at most two points and hence we may assume that E contains at least three points. In this case E is a perfect set and for any z the orbit of z,

$$\Pi z = \{g(z) : g \in G\}$$

is dense in E. Since $g(E) = E$ for all $g \in G$ Lemma 2(ii) implies that

$$D_z(E) = D_{g(z)}(E)$$

and so $D_z(E)$ is constant on each orbit. It follows immediately from these remarks that for $z \in E$, $D_z(E) = D(E)$.

COROLLARY. *Let T be a non-singular bilinear transformation. Then $G^* = TGT^{-1}$ is a properly discontinuous group with singular set $E^* = T(E)$ and $D(E) = D(E^*)$.*

Proof. It is well known that $E^* = T(E)$ and Lemma 2(ii) implies that for $z \in E$,

$$D_z(E) = D_{T(z)}(E^*).$$

Theorem 18 implies that this is equivalent to $D(E) = D(E^*)$ as required.

IMPERIAL COLLEGE, LONDON UNIVERSITY
AND
UNIVERSITY OF MARYLAND.

REFERENCES.

[1] T. Akaza, "Poincaré theta series and singular sets of Schottky groups," *Nagoya Mathematical Journal*, vol. 24 (1964), pp. 43-65.

[2] A. F. Beardon, "On the Hausdorff dimension of general Cantor sets," *Proceedings of the Cambridge Philosophical Society*, to be published.

[3] L. R. Ford, *Automorphic functions*, 2nd ed., Chelsea Publishing Company, New York, 1951.

[4] F. Hausdorff, "Dimension und äusseres Mass," *Mathematische Annalen*, vol. 79 (1918), pp. 157-179.

[5] J. Lehner, *Discontinuous groups and automorphic functions*, Mathematical Survey 7, American Mathematical Society, Rhode Island, 1964.
[6] P. J. Myrberg, "Die Kapazität der singulären Menge der linearen Gruppe," *Annales Academiae Scientiarum Fennicae*, Ser. A, 10 (1941), pp. 19.
[7] L. Sario, "Über Riemannsche Flächen mit hebbarem Rand," *Annales Academiae Scientiarum Fennicae*, vol. 50 (1948), pp. 1-79.
[8] S. J. Taylor, "On the connexion between Hausdorff measures and generalized capacity," *Proceedings of the Cambridge Philosophical Society*, vol. 57 (1961), pp. 524-531.
[9] M. Tsuji, *Potential theory in modern function theory*, Maruzen Co., Tokyo, 1959.
[10] ———, "On the capacity of general Cantor sets," *Journal of the Mathematical Society of Japan*, vol. 5 (1953), pp. 235-252.

INTEGRAL REPRESENTATIONS OF HOLOMORPHIC FUNCTIONS.

CALCU UNIVER

By GIULIANO SORANI.

1. Introduction. The integral representation of a holomorphic function of several complex variables is usually given by an integral over a $(2n-1)$-dimensional cycle. The kernel-form, K^{2n-1}, in this integral is a differential form of type $(n, n-1)$ and, as in the case of Cauchy's formula, does not depend on the domain of integration. The first proofs of this formula were given by S. Bochner [1] and E. Martinelli [4].

In this paper we shall prove a general formula giving the representation of a holomorphic function by an integral taken over a cycle of dimension $2n-p-1$, $(p=0,1,\cdots,n-1)$. Moreover, we shall show that the kernel-form K^{2n-p-1}, which appears in our formula, is a differential form of type $(n, n-p-1)$ which is connected to K^{2n-1} by a residue formula.

We remark that a general formula with a different kernel was proved by E. Martinelli [5] in 1953.

2. Homology of $X^{n,p}$. Let C^n be the complex space of complex dimension n, let $z=(z_1,\cdots,z_n)\in C^n$.

We shall denote by C_j^p, $(0\leqq p\leqq n-1, j=n-p,\cdots,n)$ *the* $p+1$ p-dimensional complex subspaces of C^n:

$$C_j^p=\{z\in C^n \mid z_1=\cdots=z_{n-p-1}=z_j=0\}.$$

Let

$$C^{n-p}=\{z\in C^n \mid z_{n-p+1}=\cdots=z_n=0\}.$$

We put

$$X_{j_1\cdots j_k}{}^{n,p}=C^n-\bigcup_{j=j_1}^{j_k} C_j^p \qquad (k=1,\cdots,p+1),$$

$$X^{n,p}=C^n-\bigcup_{j=n-p}^{n} C_j^p=X_{n-p\cdots n}{}^{n,p}.$$

LEMMA 1. *If* $p\geqq 1$, $n\geqq 2$, *then,*

$$H_i(X_{n-p}{}^{n,p},Z)=0 \text{ for } i\geqq 2n-p-1,$$

$$H_i(X_{n-p+1\cdots n}{}^{n,p},Z)\cong H_i(X^{n-1,p-1},Z).$$

Received June 7, 1965.

Proof. Since $X_{n-p}{}^{n,p}$ is contractible along the directions $z_{n-p+1},\cdots,z_n$ to $\mathbf{C}^{n-p}-\{0\}$, then

$$H_i(X_{n-p}{}^{n,p},\mathbf{Z})\cong H_i(\mathbf{C}^{n-p}-\{0\},\mathbf{Z})=0,$$

for $i\geqq 2n-2p$ and thus for $i\geqq 2n-p-1$ since $p\geqq 1$.

Next since $X_{n-p+1\cdots n}{}^{n,p}$ is contractible along the direction z_{n-p} to $X^{n-1,p-1}$, all homology groups of these two spaces are isomorphic and this proves the lemma.

THEOREM 1. $H_{2n-p-1}(X^{n,p},\mathbf{Z})\cong\mathbf{Z}$. *A generator of this homology group may be represented by the cycle*

$$\Gamma_{2n-p-1}=\tilde{S}\times S^{(n-p+1)}\times\cdots\times S^{(n)}$$

where $\tilde{S}\subset\mathbf{C}^{n-p}$ *denotes the sphere of the equation* $\sum_{i=1}^{n-p}z_i\bar{z}_i=1$ *and* $S^{(k)}$ *denotes the circle of the equations* $z_k\bar{z}_k=1,\ z_i=0,\ (i=1,\cdots,\hat{k},\cdots,n)$.

Proof. By induction on p we shall assume that, for every n,

$$H_i(X^{n,p},\mathbf{Z})=0 \text{ for } i>2n-p-1$$

$$H_{2n-p-1}(X^{n,p},\mathbf{Z})\cong\mathbf{Z},$$

and a generator of the latter group is represented by the cycle Γ_{2n-p-1}.

If $p=0$ the assumption is true since $X^{n,0}=\mathbf{C}^n-\{0\}$ is contractible to the cycle Γ_{2n-1}.

We now suppose $p\geqq 1$. Since

$$X_{n-p+1\cdots n}{}^{n,p}\cap X_{n-p}{}^{n,p}=X^{n,p}$$

$$X_{n-p+1\cdots n}{}^{n,p}\cup X_{n-p}{}^{n,p}=X^{n,p-1}$$

we may write the Mayer-Vietoris sequence of the pair $X_{n-p+1\cdots n}{}^{n,p}$ and $X_{n-p}{}^{n,p}$ as:

$$\cdots\to H_{i+1}(X_{n-p+1\cdots n}{}^{n,p},\mathbf{Z})\oplus H_{i+1}(X_{n-p}{}^{n,p},\mathbf{Z})\to H_{i+1}(X^{n,p-1},\mathbf{Z})$$
$$\to H_i(X^{n,p},\mathbf{Z})\to H_i(X_{n-p+1\cdots n}{}^{n,p},\mathbf{Z})\oplus H_i(X_{n-p}{}^{n,p},\mathbf{Z})\to\cdots.$$

Then the proof of the theorem follows by Lemma 1 and by the inductive hypothesis.

3. Kernel forms.

Our aim is to prove the following

THEOREM 2. *Let* U *be an open set in* $\mathbf{C}^n$ *containing the origin. Let* $f(z)$ *be a holomorphic function on* U. *Then*

$$(1) \qquad Nf(0) = \frac{1}{(2\pi i)^n} \int_{\gamma_{2n-p-1}} f(z)\psi^{0,n-p-1} \wedge dz_1 \wedge \cdots \wedge dz_n$$

where

$\gamma_{2n-p-1} \in H_{2n-p-1}(X^{n,p}, \mathbf{Z})$ *and*

$$(2) \quad \psi^{0,n-p-1} = \frac{\epsilon(n-p-1)! \sum_{1 \leqq \alpha_1 < \cdots < \alpha_{p+1} \leqq n} (-1)^{\alpha_1 - 1} \bar{z}_{\alpha_1} \cdots \bar{z}_{\alpha_{p+1}} d\bar{z}_1 \wedge \cdots [\alpha_1 \cdots \alpha_{p+1}] \cdots \wedge d\bar{z}_n}{(z_1\bar{z}_1 + z_2\bar{z}_2 + \cdots + z_{n-p-1}\bar{z}_{n-p-1} + z_{n-p}\bar{z}_{n-p} \cdots z_n\bar{z}_n)^{n-p}}$$

with:

$$N \in \mathbf{Z}$$
$$n - p + i - 1 \leqq \alpha_{i+1} \leqq n - p + i \qquad (i = 1, \cdots, p)$$
$$1 \leqq \alpha_1 \leqq n - p$$
$$\epsilon = (-1)^{\frac{(n-p)(n-p-1)}{2}}$$

and where $[\alpha_1 \cdots \alpha_{p+1}]$ *denotes that the differentials* $d\bar{z}_{\alpha_1}, \cdots, d\bar{z}_{\alpha_{p+1}}$ *are missing.*

We begin with a preliminary construction and a lemma, which will be needed for the proof of the theorem.

We start by constructing the differential form $\psi^{0,n-p-1}$.

α) Let $\mathfrak{U} = \{U_i\}_{1 \leqq i \leqq n-p}$ be the open covering of $X^{n,p}$ with holomorphically complete open sets given by

$$U_1 = \{z \in \mathbf{C}^n \mid z_1 \neq 0\}$$
$$\cdots\cdots\cdots\cdots\cdots$$
$$U_{n-p-1} = \{z \in \mathbf{C}^n \mid z_{n-p-1} \neq 0\}$$
$$U_{n-p} = \{z \in \mathbf{C}^n \mid z_{n-p} z_{n-p+1} \cdots z_n \neq 0\}.$$

Let $\mathcal{O}$ be the sheaf of germs of holomorphic functions on $\mathbf{C}^n$, let $A^{r,s}$ be the sheaf of germs of differential forms of type (r, s), C^∞ on $\mathbf{C}^n$.

We consider the sequence of sheaves and homomorphisms

$$(3) \qquad 0 \to \mathcal{O} \xrightarrow{i} A^{0,0} \xrightarrow{\bar{\partial}} A^{0,1} \xrightarrow{\bar{\partial}} \cdots \xrightarrow{\bar{\partial}} A^{0,n} \to 0,$$

which is an acyclic resolution of $\mathcal{O}$.

Let

$$C^q(\mathfrak{U}, \mathcal{O}) = \prod_{i_1 \cdots i_{q+1}} \Gamma(U_{i_1} \cap \cdots \cap U_{i_{q+1}}, \mathcal{O})$$

be the group of q-cochains on the covering $\mathfrak{U}$ with values in the sheaf $\mathcal{O}$. We shall denote by $\tilde{C}^q(\mathfrak{U}, \mathcal{O}) \subset C^q(\mathfrak{U}, \mathcal{O})$ the group of alternating q-cochains defined, with obvious notations, by

$$f_{i_1\cdots i_{q+1}} = \operatorname{sgn}\begin{pmatrix} i_1 \cdots i_{q+1} \\ j_1 \cdots j_{q+1} \end{pmatrix} f_{j_1\cdots j_{q+1}}.$$

We consider the homomorphism

$$\delta\colon \tilde{C}^q(\mathfrak{U}, \mathcal{O}) \to \tilde{C}^{q+1}(\mathfrak{U}, \mathcal{O})$$

defined by

$$(\delta f_{i_1\cdots i_{q+1}})_{i_1\cdots i_{q+2}} = \sum_{h=1}^{q+2} (-1)^{h-1} f_{i_1\cdots \hat{i}_h \cdots i_{q+2}}$$

and put

$$\operatorname{Ker} \delta = \tilde{Z}^q(\mathfrak{U}, \mathcal{O}).$$

The groups $C^q(\mathfrak{U}, A^{0,k})$, $\tilde{C}^q(\mathfrak{U}, A^{0,k})$ of q-cochains on $\mathfrak{U}$ with values in the sheaves $A^{0,k}$ $(0 \leqq k \leqq n)$ are defined in an analogous way.

In particular,

$$\tilde{C}^{n-p-1}(\mathfrak{U}, \mathcal{O}) = \Gamma(U_1 \cap \cdots \cap U_{n-p}, \mathcal{O}),$$
$$\tilde{C}^{n-p}(\mathfrak{U}, \mathcal{O}) = 0.$$

The latter means that the homomorphism

$$\delta\colon \tilde{C}^{n-p-1}(\mathfrak{U}, \mathcal{O}) \to \tilde{C}^{n-p}(\mathfrak{U}, \mathcal{O})$$

is the zero homomorphism. It follows that

$$\tilde{C}^{n-p-1}(\mathfrak{U}, \mathcal{O}) = \tilde{Z}^{n-p-1}(\mathfrak{U}, \mathcal{O});$$

that is, every $(n-p-1)$-cochain is a cocycle.

Let $\gamma \in \tilde{Z}^{n-p-1}(\mathfrak{U}, \mathcal{O})$. From the injection $\mathcal{O} \xrightarrow{i} A^{0,0}$, we have

$$i(\gamma) \in \Gamma(U_1 \cap \cdots \cap U_{n-p}, A^{0,0}) = \tilde{C}^{n-p-1}(\mathfrak{U}, A^{0,0}).$$

Next since γ is a cocycle, it follows that $i(\gamma) \in \tilde{Z}^{n-p-1}(\mathfrak{U}, A^{0,0})$. Moreover, since the resolution (3) is acyclic, every cocycle with values in $A^{0,0}$ is a coboundary. Therefore

$$i(\gamma) = \delta g^{0,0}{}_{(1)},$$

where

$$g^{0,0}{}_{(1)} \in \tilde{C}^{n-p-2}(\mathfrak{U}, A^{0,0}) = \prod_j \Gamma(U_1 \cap \cdots \cap \hat{U}_j \cap \cdots \cap U_{n-p}, A^{0,0}).$$

Now if

$$g^{0,0}{}_{(1),1\cdots\hat{j}\cdots n-p} \in \Gamma(U_1 \cap \cdots \cap \hat{U}_j \cap \cdots \cap U_{n-p}, A^{0,0}),$$

then

$$(\delta g^{0,0}{}_{(1),1\cdots\hat{j}\cdots n-p})_{1\cdots n-p} = \sum_{h=1}^{n-p} (-1)^{h-1} g^{0,0}{}_{(1),1\cdots\hat{h}\cdots n-p}$$

$$= \sum_{h=1}^{n-p} (-1)^{h-1} g^{0,0}{}_{(1),h},$$

where $g^{0,0}{}_{(1),h}$ is a C^∞ function defined on $U_1 \cap \cdots \cap \hat{U}_h \cap \cdots \cap U_{n-p}$. Moreover,

$$\sum_{h=1}^{n-p} (-1)^{h-1} \bar{\partial} g^{0,0}{}_{(1),h} = 0;$$

thus from the map $A^{0,0} \xrightarrow{\bar{\partial}} A^{0,1}$, it follows that

$$\bar{\partial} g^{0,0}{}_{(1),h} \in \bar{Z}^{n-p-2}(\mathcal{U}, A^{0,1}).$$

Again since the resolution (3) is acyclic, every cocycle with values in $A^{0,1}$ is a coboundary; then if we put $\bar{\partial} g^{0,0}{}_{(1),h} = g^{0,1}{}_{(1)}$, we have

$$g^{0,1}{}_{(1)} = \delta g^{0,1}{}_{(2)},$$

where $g^{0,1}{}_{(2)} \in \bar{C}^{n-p-3}(\mathcal{U}, A^{0,1})$. Similarly

$$\bar{\partial} g^{0,1}{}_{(2),k} = g^{0,2}{}_{(2)} \in \bar{Z}^{n-p-3}(\mathcal{U}, A^{0,2}) \text{ and } g^{0,2}{}_{(2)} = \delta g^{0,2}{}_{(3)}.$$

Proceeding in the same way we shall determine an element

$$g^{0,n-p-1}{}_{(n-p-1)} \in \bar{Z}^0(\mathcal{U}, A^{0,n-p-1}).$$

β) We consider the function $1/z_1 \cdots z_n \in Z^{n-p-1}(\mathcal{U}, \mathcal{O})$. This function is a cocycle which is not a coboundary on $U_1 \cap \cdots \cap U_{n-p}$. Therefore

$$i\left(\frac{1}{z_1 \cdots z_n}\right) = \frac{a_1}{z_2 \cdots z_n} + \frac{a_2}{z_1 \hat{z}_2 \cdots z_n} + \cdots + \frac{a_{n-p}}{z_1 \cdots z_{n-p-1}}.$$

It follows that

$$(4) \begin{cases} a_1 = \bar{z}_1 / (z_1\bar{z}_1 + \cdots + z_{n-p-1}\bar{z}_{n-p-1} + z_{n-p}\bar{z}_{n-p} \cdots z_n\bar{z}_n) \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_h = (-1)^{h-1}\bar{z}_h / (z_1\bar{z}_1 + \cdots + z_{n-p-1}\bar{z}_{n-p-1} + z_{n-p}\bar{z}_{n-p} \cdots z_n\bar{z}_n) \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{n-p} = (-1)^{n-p-1}\bar{z}_{n-p} \cdots \bar{z}_n / (z_1\bar{z}_1 + \cdots + z_{n-p-1}\bar{z}_{n-p-1} + z_{n-p}\bar{z}_{n-p} \cdots z_n\bar{z}_n). \end{cases}$$

Thus we have

$$g_h = (-1)^{h-1}\frac{a_h}{z_1\cdots\hat{z}_h\cdots z_n} \text{ if } h \leqq n-p-1$$

$$g_{n-p} = (-1)^{n-p-1}\frac{a_{n-p}}{z_1\cdots z_{n-p-1}}$$

with a_h $(h=1,\cdots,n-p)$ given by (4).

We now consider g_{n-p}, on $U_1\cap\cdots\cap U_{n-p-1}$, and compute $\bar{\partial}g_{n-p}$.

$$\begin{aligned}\bar{\partial}g_{n-p} &= (-1)^{n-p-1}\bar{\partial}\left(\frac{a_{n-p}}{z_1\cdots z_{n-p-1}}\right)\\ &= (-1)^{n-p-1}\bar{\partial}\left\{\bar{z}_{n-p}\cdots\bar{z}_n\frac{(z_1\bar{z}_1+\cdots+z_{n-p-1}\bar{z}_{n-p-1}+z_{n-p}\bar{z}_{n-p}\cdots z_n\bar{z}_n)^{-1}}{z_1\cdots z_{n-p-1}}\right\}\\ &= (-1)^{n-p-1}\frac{1}{z_1\cdots z_{n-p-1}}\{(\bar{z}_{n-p+1}\cdots\bar{z}_n d\bar{z}_{n-p}+\bar{z}_{n-p}\bar{z}_{n-p+2}\cdots\bar{z}_n d\bar{z}_{n-p+1}+\cdots\\ &\qquad +\bar{z}_{n-p}\cdots\bar{z}_{n-1}d\bar{z}_n)(z_1\bar{z}_1+\cdots+z_{n-p-1}\bar{z}_{n-p-1}+z_{n-p}\bar{z}_{n-p}\cdots z_n\bar{z}_n)\\ &\qquad -\bar{z}_{n-p}\cdots\bar{z}_n(z_1 d\bar{z}_1+\cdots+z_{n-p-1}d\bar{z}_{n-p-1}+z_{n-p}z_{n-p+1}\bar{z}_{n-p+1}\cdots\\ &\qquad\qquad \cdots z_n\bar{z}_n d\bar{z}_{n-p}+\cdots\\ &\quad + z_{n-p}\bar{z}_{n-p}\cdots z_{n-1}\bar{z}_{n-1}z_n d\bar{z}_n))\}(z_1\bar{z}_1+\cdots+z_{n-p-1}\bar{z}_{n-p-1}+z_{n-p}\bar{z}_{n-p}\cdots z_n\bar{z}_n)^{-2}\\ &= (-1)^{n-p-1}\frac{(z_1\bar{z}_1+\cdots+z_{n-p-1}\bar{z}_{n-p-1}+z_{n-p}\bar{z}_{n-p}\cdots z_n\bar{z}_n)^{-2}}{z_1\cdots z_{n-p-1}}\{(\bar{z}_{n-p+1}\cdots\bar{z}_n d\bar{z}_{n-p}\\ &\qquad +\cdots+\bar{z}_{n-p}\cdots\bar{z}_{n-1}d\bar{z}_n)(z_1\bar{z}_1+\cdots+z_{n-p-1}\bar{z}_{n-p-1})\\ &\qquad -\bar{z}_{n-p}\cdots\bar{z}_n(z_1 d\bar{z}_1+\cdots+z_{n-p-1}d\bar{z}_{n-p-1})\}.\end{aligned}$$

In the latter expression we consider that term which is defined on $U_1\cap\cdots\cap U_{n-p-2}$, namely:

$$\begin{aligned}\psi^{0,1} &= \frac{a_{n-p,n-p-1}}{z_1\cdots z_{n-p-2}}\\ &= (-1)^{(n-p-1)+(n-p-2)}\ \frac{(z_1\bar{z}_1+\cdots+z_{n-p-1}\bar{z}_{n-p-1}+z_{n-p}\bar{z}_{n-p}\cdots z_n\bar{z}_n)^{-2}}{z_1\cdots z_{n-p-2}}\\ &\qquad (\bar{z}_{n-p-1}\bar{z}_{n-p+1}\cdots\bar{z}_n d\bar{z}_{n-p}+\cdots+\bar{z}_{n-p-1}\cdots\bar{z}_{n-1}d\bar{z}_n-\bar{z}_{n-p}\cdots\bar{z}_n d\bar{z}_{n-p-1}),\end{aligned}$$

and we compute $\bar{\partial}\psi^{0,1}$. The computation gives

$$\begin{aligned}\bar{\partial}\psi^{0,1} &= (-1)^{(n-p-1)+(n-p-2)}\frac{2(z_1\bar{z}_1+\cdots+z_{n-p-1}\bar{z}_{n-p-1}+z_{n-p}\bar{z}_{n-p}\cdots z_n\bar{z}_n)^{-3}}{z_1\cdots z_{n-p-2}}\\ &\quad \{(\bar{z}_{n-p+1}\cdots\bar{z}_n d\bar{z}_{n-p-1}\wedge d\bar{z}_{n-p}+\cdots+\bar{z}_{n-p}\cdots\bar{z}_{n-1}d\bar{z}_{n-p-1}\wedge d\bar{z}_n)(z_1\bar{z}_1+\cdots\\ &\qquad +z_{n-p-2}\bar{z}_{n-p-2})+(\bar{z}_{n-p-1}\bar{z}_{n-p+1}\cdots\bar{z}_n d\bar{z}_{n-p}+\cdots+\bar{z}_{n-p-1}\cdots\end{aligned}$$

$$\bar{z}_{n-1}d\bar{z}_n - \bar{z}_{n-p}\cdots\bar{z}_n d\bar{z}_{n-p-1}) \wedge (z_1 d\bar{z}_1 + \cdots + z_{n-p-2}d\bar{z}_{n-p-2})\}.$$

Similarly we consider, on $U_1 \cap \cdots \cap U_{n-p-3}$,

$$\psi^{0,2} = \frac{a_{n-p,n-p-1,n-p-2}}{z_1\cdots z_{n-p-3}}$$

$$= (-1)^{(n-p-1)+(n-p-2)+(n-p-3)} \frac{2(z_1\bar{z}_1 + \cdots + z_{n-p-1}\bar{z}_{n-p-1} + z_{n-p}\bar{z}_{n-p}\cdots z_n\bar{z}_n)^{-3}}{z_1\cdots z_{n-p-3}}$$

$$(\bar{z}_{n-p-2}\bar{z}_{n-p+1}\cdots\bar{z}_n d\bar{z}_{n-p-1} \wedge d\bar{z}_{n-p} + \cdots + \bar{z}_{n-p-2}\bar{z}_{n-p}\cdots\bar{z}_{n-1}d\bar{z}_{n-p-1} \wedge d\bar{z}_n$$

$$+ \bar{z}_{n-p-1}\bar{z}_{n-p+1}\cdots\bar{z}_n d\bar{z}_{n-p} \wedge d\bar{z}_{n-p-2} + \cdots + \bar{z}_{n-p-1}\cdots\bar{z}_{n-1}d\bar{z}_n \wedge d\bar{z}_{n-p-2}$$

$$- \bar{z}_{n-p}\cdots\bar{z}_n d\bar{z}_{n-p-1} \wedge d\bar{z}_{n-p-2}).$$

Iterating this procedure, we get after $n-p-3$ additional steps the differential form $\psi^{0,n-p-1}$ shown in (2).

We put

$$K^{2n-p-1} = \psi^{0,n-p-1} \wedge dz_1 \wedge \cdots \wedge dz_n.$$

4. Residues of kernel-forms. Throughout this section we shall apply the methods developed by J. Léray [2]. Let $\boldsymbol{C}_k{}^{n-1} = \{z \in \boldsymbol{C}^n \mid z_k = 0\}$. We put

$$T^0 = \boldsymbol{C}^n - \boldsymbol{C}_{n-p}{}^p \cup \cdots \cup \boldsymbol{C}_n{}^p = X^{n,p},$$

$$T^1 = \boldsymbol{C}_{n-p+1}{}^{n-1} - \boldsymbol{C}_{n-p+1}{}^p \cap (\boldsymbol{C}_{n-p}{}^p \cup \boldsymbol{C}_{n-p+2}{}^p \cup \cdots \cup \boldsymbol{C}_n{}^p),$$

$$T^i = \boldsymbol{C}_{n-p+1}{}^{n-1} \cap \cdots \cap \boldsymbol{C}_{n-p+i}{}^{n-1} - \boldsymbol{C}_{n-p+1}{}^p \cap \cdots \cap \boldsymbol{C}_{n-p+i}{}^p$$

$$\cap (\boldsymbol{C}_{n-p}{}^p \cup \boldsymbol{C}_{n-p+i+1}{}^p \cup \cdots \cup \boldsymbol{C}_n{}^p)$$

$$T^p = \boldsymbol{C}_{n-p+1}{}^{n-1} \cap \cdots \cap \boldsymbol{C}_n{}^{n-1} - \boldsymbol{C}_{n-p+1}{}^p \cap \cdots \cap \boldsymbol{C}_n{}^p \cap \boldsymbol{C}_{n-p}{}^p.$$

We consider the homomorphisms *cobord composé* and *résidu composé*:

$$\delta^p : H_*(T^p) \to H_*(T^0),$$

$$\mathrm{res}^p : H^*(T^0) \to H^*(T^p).$$

Since $T^p = \boldsymbol{C}^{n-p} - \{0\}$, we have

$$\delta^p : H_i(\boldsymbol{C}^{n-p} - \{0\}) \to H_{i+p}(\boldsymbol{C}^n - \bigcup_j \boldsymbol{C}_j{}^p)$$

$$\mathrm{res}^p : H^{i+p}(\boldsymbol{C}^n - \bigcup_j \boldsymbol{C}_j{}^p) \to H^i(\boldsymbol{C}^{n-p} - \{0\}).$$

Let ϕ^{2n-p-1} be a differential form which represents a cohomology class of $H^{2n-p-1}(\boldsymbol{C}^n - \bigcup_j \boldsymbol{C}_j{}^p)$, let $\gamma_{2n-2p-1}$ be a cycle which represents the unique homology class of $H_{2n-2p-1}(\mathbf{C}^{n-p}-\{0\})$; then the residue formula $\langle\phi,\delta^p\gamma\rangle = \langle \operatorname{res}^p\phi,\gamma\rangle$ gives

$$\int_{\delta^p\gamma}\phi^{2n-p-1}=(2\pi i)^p\int_\gamma \operatorname{res}^p\phi^{2n-p-1}. \tag{5}$$

LEMMA 2. *Let* $K^{2n-p-1}=\psi^{0,n-p-1}\wedge dz_1\wedge\cdots\wedge dz_n$. *Then*

$$\operatorname{res}^p K^{2n-p-1}\,|_p=K^{2(n-p)-1},$$

where $K^{2n-p-1}\,|_p$ *denotes the restriction of* K^{2n-p-1} *to* $S^{(n-p+1)}\times\cdots\times S^{(n)}$, res^p *denotes the iterated residue with respect to* $\mathbf{C}^{n-p}$. *Putting* $m=n-p$, *then* $K^{2m-1}=\psi^{0,m-1}\wedge dz_1\wedge\cdots\wedge dz_m$, *where* $\psi^{0,m-1}$ *is given by* (2) *with* $n=m$ *and* $p=0$.

Proof. On Γ_{2n-p-1}, we have $\bar z_k=\dfrac{1}{z_k}$, $d\bar z_k=-\dfrac{\bar z_k}{z_k}dz_k$, $(k=n-p+1,\cdots,n)$. Hence

$$K^{2n-p-1}\,|_p=(-1)^{\frac{m(m-1)}{2}}(m-1)!\sum_{\alpha=1}^{m}(-1)^{\alpha-1}\bar z_\alpha(z_1\bar z_1+\cdots+z_m\bar z_m)^{-m}$$
$$d\bar z_1\wedge\cdots\wedge d\hat{\bar z}_\alpha\wedge\cdots\wedge d\bar z_m\wedge dz_1\wedge\cdots\wedge dz_m\wedge\frac{dz_{m+1}}{z_{m+1}}\wedge\cdots\wedge\frac{dz_n}{z_n}.$$

Now the proof of the lemma follows from the definition of iterated residue.

COROLLARY. *We have*

$$\int_{\Gamma_{2n-p-1}}K^{2n-p-1}=(2\pi i)^p\int_{\bar S}K^{2m-1}.$$

Since $\Gamma_{2n-p-1}=\delta^p\bar S$, the proof follows from the residue formula (5).

5. Proof of Theorem 2. We observe that any cycle γ_{2n-p-1} which is not a boundary in $X^{n,p}$ is an integral multiple of Γ_{2n-p-1}, so it suffices to prove (1) with $\gamma_{2n-p-1}=\Gamma_{2n-p-1}$.

By a suitable choice of the unit length we may assume without loss that $\Gamma_{2n-p-1}\subset U$.

First of all we observe that from the corollary of Lemma 2, it follows that

$$\int_{\Gamma_{2n-p-1}}f(z)K^{2n-p-1}=(2\pi i)^p\int_{\bar S}f(z)K^{2m-1},$$

where $f(z)$ in the right-hand side is a function of $(z_1,\cdots,z_m)$ only. Thus we need only prove (1) for $p=0$; that is, we must show that, up to a sign,

$$\int_{S_{2n-1}} f(z) K^{2n-1} = (2\pi i)^n f(0), \tag{6}$$

where S_{2n-1} denotes the sphere in $\boldsymbol{C}^n$ of the equation $\sum_{i=1}^{n} z_i \bar{z}_i = 1$.

But we are reduced to the classical $(2n-1)$-dimensional case mentioned in the intrdouction. The proof of this result is standard, and can be carried out with the aid of Stokes' theorem. However, for completeness, we give here a sketch of this proof.

First of all, as is easy to check, since $d(f(z)K^{2n-1}) = 0$, the integral in (6) does not depend on the radius of the sphere S_{2n-1}, so that its value equals $f(0) \int_{S_{2n-1}} K^{2n-1}$.

Next by Stokes theorem we have

$$\int_{S_{2n-1}} K^{2n-1} = (-1)^n n! \int_{B_{2n}} dz_1 \wedge \cdots \wedge dz_n \wedge d\bar{z}_1 \wedge \cdots \wedge d\bar{z}_n,$$

where B_{2n} denotes the unit ball in $\boldsymbol{C}^n$.

Now putting $z_k = x_k + iy_k$ $(k = 1, \cdots, n)$ and taking

$$dz_1 \wedge \cdots \wedge dz_n \wedge d\bar{z}_1 \wedge \cdots \wedge d\bar{z}_n$$
$$= \frac{2^n}{i^n} dx_1 \wedge \cdots \wedge dx_n \wedge dy_1 \wedge \cdots \wedge dy_n,$$

we have

$$\int_{S_{2n-1}} K^{2n-1} = i^n n! \, 2^n \int_{B_{2n}} dx_1 \wedge \cdots \wedge dx_n \wedge dy_1 \wedge \cdots \wedge dy_n = (2\pi i)^n$$

since the volume of B_{2n} is equal to $\dfrac{\pi^n}{n!}$.

Remark. There are some slight differences between our $(2n-1)$-dimensional kernel and those of Bochner and Martinelli. The factor $(-1)^{\frac{n(n-2)}{2}}$ which appears in our kernel is due to the choice of orientation. Putting $\rho^2 = \sum_{i=1}^{n} z_i \bar{z}_i$, the Martinelli kernel may be written as $\dfrac{2^{n-1} i^n}{1-n} * \partial \rho^{2(1-n)}$ so that the factor $(n-1)!$ does not appear in his kernel. However, this factor of $(n-1)!$ is included in our kernel, as in that of Bochner, and consequently it does not appear in formula (6).

NORTHEASTERN UNIVERSITY,
BOSTON, MASSACHUSETTS.

REFERENCES.

[1] S. Bochner, "Analytic and meromorphic continuation by means of Green's formula," *Annals of Mathematics*, vol. 44 (1943), p. 652.

[2] J. Léray, "Le calcul différentiel et intégral sur une variété analytique complexe," *Bulletin de la Société Mathématique de France*, vol. 87 (1959), pp. 81-180.

[3] E. Martinelli, "La formula di Cauchy per le funzioni analitiche di due variabili complesse," *Rendiconti dell' Accademia Nazionale dei Lincei*, vol. XXV, serie 6 (1937), pp. 33-36.

[4] E. Martinelli, "Alcuni teoremi integrali per le funzioni di più variabili complesse," *Memorie dell' Accademia d'Italia* (1938).

[5] ———, "Sulle estensioni della formula integrale di Cauchy alle funzioni analitiche di più variabili complese," *Annali di Matematica pura e applicata*, serie 4^a, (1953), pp. 277-347.

LIFTING OF TWO-DIMENSIONAL SINGULARITIES TO CHARACTERISTIC ZERO.

By Michael Artin.*

1. Introduction. Let k be an algebraically closed field of characteristic $p \neq 0$, and let $\bar{A}$ be a complete normal local ring of dimension 2 over k. *Mumford* has posed the problem of lifting the ring $\bar{A}$ in an "equisingular" way to characteristic zero. In this paper we give a solution of a slightly weaker problem: We lift, not $\bar{A}$, but a subring $\bar{A}' \subset \bar{A}$ whose normalization is $\bar{A}$ and such that the conductor of $\bar{A}'$ in $\bar{A}$ is $\mathfrak{m}'$-primary ($\mathfrak{m}'$ is the maximal ideal of $\bar{A}'$). This means that the map $\operatorname{Spec} \bar{A} \rightarrow \operatorname{Spec} \bar{A}'$ is an isomorphism outside of the closed point, and so the "topological" properties of the spectra are the same. The solution is therefore satisfactory from many points of view. Because of the primitive state of the theory of deformations of singularities, we have left the result as it came, without entering into a discusssion of the definition of equisingularity. The properties of the lifting (cf. Section 2) seem to us to be about as good as could be wished. The proof uses routine infinitesimal methods. In Section 4, we describe which extensions of $\bar{A}'$, unramified outside of the closed point, lift to characteristic zero, and in the last section we discuss an example of *Mumford* of a wildly ramified local extension.

2. The Theorem. Let R be a complete discrete valuation ring with algebraically closed residue field k (of any characteristic), and let t be a local parameter for R. Let $\bar{A}$ be a complete normal ring of dimension 2 over k, and $\bar{X} \rightarrow \operatorname{Spec} \bar{A}$ a proper birational map with $\bar{X}$ nonsingular (i. e., a resolution of the singularity of $\operatorname{Spec} \bar{A}$ [1, 13]) such that the closed fibre of $\bar{X}/\operatorname{Spec} \bar{A}$ is set-theoretically the union $\bar{C} = \bigcup_i \bar{C}^i$ of nonsingular curves $\bar{C}^i$ $(i = 1, \cdots, r)$ with only normal crossings at the points P^j $(j = 1, \cdots, s)$.

Theorem (2.1). *There exists a complete normal ring A of dimension 3 containing R, together with a homomorphism $A \rightarrow R$ inducing the identity on R (a section $S \colon \operatorname{Spec} R \rightarrow \operatorname{Spec} A$) and a birational proper map $f \colon X \rightarrow \operatorname{Spec} A$, with X nonsingular (i. e., a resolution), having the following properties:*

Received July 29, 1965.

* This research was partially supported by Grant GP-4362 of the National Science Foundation.

(a) *f is an isomorphism outside of the section S, and the reduced inverse image of S in X is the union C of curves C^i smooth over R $(i=1,\cdots,r)$ with normal crossings along the sections P^j $(j=1,\cdots,s)$ of $X/\operatorname{Spec} R$.*

(b) *Let $\otimes_R k$ denote the change of base $\operatorname{Spec} k \to \operatorname{Spec} R$. Then $X \otimes_R k = \bar{X}$, $C^i \otimes_R k = \bar{C}^i$, and $P^j \otimes_R k = \bar{P}^j$.*

(c) *$A \otimes_R k = \bar{A}'$ is a subring of $\bar{A}$ having $\bar{A}$ as its normalization, and the conductor of $\bar{A}'$ in $\bar{A}$ is $\mathfrak{m}'$-primary ($\mathfrak{m}'$ the maximal ideal of $\bar{A}'$).*

(d) *The map $X \to \operatorname{Spec} R$ is geometrically regular, i. e., has regular geometric fibres.*

To make the construction, we first lift each curve $\bar{C}^i$ to a smooth and projective curve C^i over $\operatorname{Spec} R$ [9, p. 182-14, Cor. 4]. For each intersection $\bar{P}$ of $\bar{C}$ and $\bar{C}^{i'}$, choose sections S^i, $S^{i'}$ on C^i and on $C^{i'}$ specializing to $\bar{P}$. Since C^i, $C^{i'}$ are smooth over $\operatorname{Spec} R$ and since R is complete, this can be done. Now as is immediately seen, there is a reduced scheme C over R whose components are the C^i and where, for each intersection $\bar{P}$, C^i and $C^{i'}$ intersect transversally along the sections S^i, $S^{i'}$. We denote the common section by P. (The regular functions on an open subset U of C are n-tuples of functions $f=(f_1,\cdots,f_r)$, $f_i \in \Gamma(U \cap C^i, \mathcal{O}_{C^i})$ such that, with the above notation, $f_i \mid S^i \cap U = f_{i'} \mid S^{i'} \cap U$.) Thus we can construct a candidate for the reduced inverse image of S in X.

Next, we have to find the formal scheme $\mathfrak{X}$ which will be the completion of X along C. A fundamental system of infinitesimal neighborhoods of $\bar{C}$ in $\bar{X}$ is given by "cycles" of the form

$$\bar{Y}_n = \sum_{i=1}^{r} n_i C^i$$

where $n=(n_1,\cdots,n_r)$ is an r-tuple of positive integers. The scheme $\bar{Y}_n$ is the one defined by the sheaf of ideals on $\bar{X}$ of functions which vanish with order at least n_i on $\bar{C}^i$. Hence the formal neighborhood of $\bar{C}$ in $\bar{X}$ is

$$\mathfrak{X} = \lim_{\substack{\rightarrow \\ n}} \bar{Y}_n .$$

The formal scheme $\mathfrak{X}$ over $\operatorname{Spec} R$ will have to induce $\mathfrak{X}$ $(\operatorname{mod} t)$, and will have to contain an ideal of definition J which defines the closed subscheme C of $\mathfrak{X}$. In addition, we require that $\mathfrak{X}$ be *regular*, i. e., that there exist a formal affine covering by opens $U_i = \operatorname{Spec} B_i$ where B_i is a noetherian admissible ring [6, Ch. 0.7] which is regular at every point of $\operatorname{Spf} B_i$.

PROPOSITION (2.2). *There exists a formal scheme $\mathfrak{X}$ having the above properties.*

We will give the proof in Section 3. Roughly speaking, the proposition follows from the fact that $\bar{C}$ has dimension 1, which assures that the obstructions to the construction of $\mathfrak{X}$ step by step vanish. For the remainder of this section we assume that we are given such a scheme.

Since each curve $\bar{C}^i$ is defined locally in $\mathfrak{X}$ by one equation (or since $\mathfrak{X}$ is regular), it is clear that the subscheme C^i of $\mathfrak{X}$ is also defined by one equation locally. For each $n = (n_1, \cdots, n_r)$ we can introduce the subscheme Y_n of $\mathfrak{X}$ defined by the ideal of functions on $\mathfrak{X}$ which vanish to order at least n_i on C^i. Since the ideal J defining C is an ideal of definition, the subschemes Y_n form a fundamental system of formal neighborhoods of C.

Let $n = (n_1, \cdots, n_r)$ be chosen such that the intersection numbers $(\bar{Y}_n \cdot \bar{C}^i)$ are negative for $i = 1, \cdots, r$. Since the intersection matrix $\| (\bar{C}^i \cdot \bar{C}^j) \|$ is negative definite [12], such an n exists with all $n_i > 0$. For $\nu \geqq 0$, put $Z_\nu = Y_m$ and $\bar{Z}_\nu = \bar{Y}_m$, where $m = ((\nu+1)n_1, \cdots, (\nu+1)n_r)$. The Z_ν are a fundamental system of formal neighborhoods of C. Consider the kernel K of the map $\mathcal{O}_{Z_\nu} \to \mathcal{O}_{Z_{\nu-1}}$. This kernel is easily seen [2, (1.8)] to be an invertible sheaf on $\mathcal{O}_{Z_0}$ isomorphic to $L_0{}^{\otimes \nu}$ where L_0 is the restriction to Z_0 of the invertible sheaf $L = \mathcal{L}(-Z_0)$ on X defined by the "Cartier divisor" $-Z_0$. Since $(\bar{Z} \cdot \bar{C}^i) < 0$, the sheaf L induces an ample invertible sheaf on each C^i relative to Spec R, and hence on each Z_ν by [6, III. 2. 6. 2].

LEMMA (2. 3). *Let* $A = \varprojlim_n H^0(C, \mathcal{O}_{Z_\nu})$. *$A$ is a complete noetherian local ring.*

Proof. Let $I_\nu \subset A$ be the kernel of the map $A \to H^0(C, \mathcal{O}_{Z_\nu})$. To show that A is noetherian, it suffices by [3, p. 45, Cor. 5] to show that $A_0 = A/I_0$ is noetherian and that I_0/I_1 is an A_0-module of finite type. Since the kernel K_ν of $\mathcal{O}_{Z_\nu} \to \mathcal{O}_{Z_{\nu-1}}$ is isomorphic to $L_0{}^{\otimes \nu}$ and L_0 is ample, $H^1(C, K_\nu) = 0$ for sufficiently large ν, and so $H^0(C, \mathcal{O}_{Z_\nu}) \to H^0(C, \mathcal{O}_{Z_{\nu-1}})$ is surjective for large ν. Therefore any element of $H^0(C, \mathcal{O}_{Z_\nu})$ lifts to A and so $A_\nu = A/I_\nu = H^0(C, \mathcal{O}_{Z_\nu})$ if ν is large enough. Now the schemes Z_ν are proper over Spec R, and so by the finiteness theorem for cohomology [6, III. 3. 2. 1], $H^0(C, \mathcal{O}_{Z_\nu})$ is a finite R-algebra, and hence is noetherian. Therefore A_ν is noetherian for large ν, and so A_0 is, too. It is easily seen that I_0/I_1 is mapped injectively to $H^0(C, K_1)$ by the map $A_1 \to H^0(C, \mathcal{O}_{Z_1})$. Since K_1 is a coherent sheaf on Z_0, $H^0(C, K_1)$ is of finite type as an R-module [6, III. 3. 2. 1], hence as an A_0-module. Finally, A is local because the closed fibre of Z_ν over Spec R is connected, hence each $H^0(C, \mathcal{O}_{Z_\nu})$ is local, and A is complete because R, and hence each A_ν, is.

LEMMA. *For a large enough integer* N, *the invertible sheaf* $L^{\otimes N}$ *is very ample on* $\mathfrak{X}$ *relative to* Spec A. *More precisely, the restrictions* $L_\nu{}^{\otimes N}$ *of* $L^{\otimes N}$ *to* Z_ν *define compatible projective embeddings of the* Z_ν *in a suitable projective space.*

Proof. Choose N large enough so that $H^1(C, L_0{}^{\otimes N}) = 0$ if $n \geqq N$ and so that $L_1{}^{\otimes N}$ is very ample on Z_1 relative to Spec A. The kernel of $L_\nu{}^{\otimes N} \rightarrow L_{\nu-1}{}^{\otimes N}$ is $L^{\otimes N} \otimes K_\nu \approx L_0{}^{(N+\nu)}$. Hence $H^1(C, L^{\otimes N} \otimes K_\nu) = 0$, and so the global sections of $L_{\nu-1}{}^{\otimes N}$ lift to $L_\nu{}^{\otimes N}$. It follows that $L_\nu{}^{\otimes N}$ is very ample on Z_ν relative to Spec A—in fact, the section of $L_\nu{}^{\otimes N}$ separate points and tangent directions because they do when restricted to Z_1, and because Z_1 has the same underlying space and the same tangent space everywhere as Z_ν.

If n is large enough so that K_ν is generated by its sections and has trivial H^1 for $\geqq n$, then it is easily seen that

$$H^0(C, L_{\nu-1}{}^{\otimes N}) \approx H^0(C, L_\nu{}^{\otimes N}) \otimes_{A_\nu} A_{\nu-1}.$$

Therefore by the Nakayama lemma, any set of sections of $H^0(C, L_\nu{}^{\otimes N})$ which induces a generating set in $H^0(C, L_n{}^{\otimes N})$ $(\nu \geqq n)$ generates $H^0(C, L_\nu{}^{\otimes N})$ as A-module. Therefore we can choose a compatible finite generating set for these modules, and thus obtain compatible embeddings in a suitable projective space $\boldsymbol{P}^n{}_A$ over Spec A, which proves the lemma.

Moreover, it follows that $\mathfrak{X}$ is a closed formal subscheme of the formal scheme $\mathfrak{P} = \underset{\rightarrow}{\lim}\, P_\nu$ where P_ν is the projective space over Spec A_ν. By Grothendieck's existence theorem [6, III. 5. 1. 8] $\mathfrak{X}$ is the formal completion of a scheme $X \subset \boldsymbol{P}^n{}_A$, and we claim that $f \colon X \rightarrow$ Spec A satisfies the conditions of the theorem.

We first prove assertion (c) of the theorem: What we need to show is that $\bar{A}/\bar{A}'$ is a finite dimensional k-module. Let ϵ_ν be the cokernel of the map

$$H^0(C, \mathcal{O}_{Z_\nu}) \rightarrow H^0(\bar{C}, \mathcal{O}_{Z_\nu}).$$

Certainly ϵ_ν has finite length because $H^0(\bar{C}, \mathcal{O}_{\bar{Z}_\nu})$ does. We have an exact sequence

$$0 \rightarrow \bar{A}'_\nu \rightarrow H^0(\bar{C}, \mathcal{O}_{\bar{Z}_\nu}) \rightarrow \epsilon_\nu \rightarrow 0,$$

and

$$\bar{A} = \underset{\leftarrow}{\lim}\, H^0(\bar{C}, \mathcal{O}_{\bar{Z}_\nu}).$$

Thus since $\underset{\leftarrow}{\lim}$ is left exact, it suffices to show that the canonical map $\epsilon_\nu \rightarrow \epsilon_{\nu-1}$ is bijective for large ν. Since the kernel K_ν of $\mathcal{O}_{Z_\nu} \rightarrow \mathcal{O}_{Z_{\nu-1}}$ has trivial cohomology for large ν as was seen above, the map

$$H^1(C, \mathcal{O}_{Z_\nu}) \to H^1(C, \mathcal{O}_{Z_{\nu-1}})$$

is bijective for large ν. The bijectivity of the map $\epsilon_\nu \to \epsilon_{\nu-1}$ now follows from the cohomology diagram deduced from the diagram

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{O}_{Z_\nu} \arrow[r, "t"] \arrow[d] & \mathcal{O}_{Z_\nu} \arrow[r] \arrow[d] & \mathcal{O}_{\bar{Z}_\nu} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathcal{O}_{Z_{\nu-1}} \arrow[r, "t"'] & \mathcal{O}_{Z_{\nu-1}} \arrow[r] & \mathcal{O}_{\bar{Z}_{\nu-1}} \arrow[r] & 0.
\end{tikzcd}$$

This shows that assertion (c) of the theorem is true.

Now since $\mathfrak{X}$ was assumed regular, X is nonsingular at every point of C, hence is nonsingular since every point of X has a specialization on C. By the theorem of holomorphic functions [6, III. 4. 1. 7]

$$\begin{aligned} H^0(X, \mathcal{O}_X) &\approx \lim_{\leftarrow} H^0(C, \mathcal{O}_X \otimes A_\nu) \\ &= \lim_{\leftarrow} H^0(C, \mathcal{O}_{Z_\nu}) = A. \end{aligned}$$

Since X is regular and connected, it follows that A is a normal local ring. Since the image of the parameter t of R in A is not zero,

$$\dim A = \dim \bar{A}' + 1 = 3.$$

Moreover since $\bar{f}\colon \bar{X} \to \operatorname{Spec} \bar{A}'$ is birational, $f\colon X \to \operatorname{Spec} A$ is generically 1-1, and therefore birational by [6, III. 4. 3. 3].

Consider the image in $\operatorname{Spec} A$ of the subscheme C of X. It is clear from the construction of C that $H^0(C, \mathcal{O}_C) \approx R$, and so that A maps surjectively to $H^0(C, \mathcal{O}_C)$. Therefore the image of C is a subscheme isomorphic to $\operatorname{Spec} R$, and this gives the desired section $S\colon \operatorname{Spec} R \to \operatorname{Spec} A$.

We still have to show that $f\colon X \to \operatorname{Spec} A$ is an isomorphism outside of S. The remaining assertions of (a) and (b) are true by construction, and (d) is easily seen.

Suppose that f is not an isomorphism at some point of X not on C. Then since A is normal, the fibre X_y has dimension > 0 for some point y of $\operatorname{Spec} A$ not on S. For reasons of dimension, the point y must have codimension 2, i. e., must be the generic point of an irreducible subset Y of $\operatorname{Spec} A$ of dimension 1, and X_y must be of dimension 1. Let D be a reduced component of $X \times_{\operatorname{Spec} A} Y$ whose fibre D_y over y had dimension 1. Then clearly D has dimension 2 and so since X is regular, D is the support of a Cartier divisor of X. Since $\bar{f}\colon \bar{X} \to \operatorname{Spec} \bar{A}'$ is regular outside of $\bar{C}$, it follows that

$Y \not\subset \operatorname{Spec} \bar{A}'$, and so $Y \cap \operatorname{Spec} \bar{A}'$ is the closed point of $\operatorname{Spec} A$. Therefore $D \cap \bar{X}$ is a closed subset of $\bar{C}$. Let $\bar{D}$ be the scheme-theoretic intersection of D with $\bar{X}$. We can view $\bar{D}$ as a Cartier divisor on $\bar{X}$, and its support is $\bar{C}$. Therefore $\bar{D}$ is of the form $\bar{D} = \sum r_i \bar{C}^i$ $(r_i \geqq 0)$. Since $\bar{D} \neq 0$, and since $\|(\bar{C}^i \cdot \bar{C}^j)\|$ is negative definite [12], $(\bar{D} \cdot \bar{C}^i) < 0$ for some i, say for $i = 1$. This implies that the restriction of the invertible sheaf $\mathcal{L}(D)$ on X to the curve $\bar{C}^1$ has negative degree, and thus is not the trivial invertible sheaf. On the other hand, y is not on S, and hence D meets C^1 only on the closed fibre $\bar{C}^1$. Therefore the restriction of $\mathcal{L}(D)$ to C^1 is trivial on the generic fibre of $C^1/\operatorname{Spec} R$, and hence is trivial since $C^1/\operatorname{Spec} R$ is a smooth curve. A fortiori, the restriction of $\mathcal{L}(D)$ to $\bar{C}^1$ is trivial, which is a contradiction, and completes the proof of the theorem.

3. Construction of the formal scheme. We want to construct the formal scheme $\mathfrak{X}$ of (2.2) by deformation theory, and so we first have to show that each infinitesimal neighborhood Y_n is determined in advance locally up to isomorphism at any point of C. We need only consider closed points $\bar{P}$ of $\bar{C}$. Suppose first that the point $\bar{P}$ is a point of intersection of two of the C^i, say of C^1 and C^2. Choose an affine neighborhood, say $\operatorname{Spec} B = U$, of $\bar{P}$ in Y_n on which the subschemes C^1, C^2 are defined each by one equation, say $C^1 = V(x)$, $C^2 = V(y)$ for $x, y \in B$. We have $x^p = y^q = 0$ where $p = n_1$, $q = n_2$, $n = (n_1, \cdots, n_r)$. Consider the map $U = \operatorname{Spec} B \to \operatorname{Spec} R[x,y]$ (which is quasi-finite). Since by assumption $\bar{Y}_n$ is an infinitesimal neighborhood of $\bar{C}$ in $\bar{X}$, the elements x, y, t generate the ideal of $\bar{P}$ in B and hence the map is unramified at $\bar{P}$. Now it is clear from the assumption of regularity on $\mathfrak{X}$ that the completion $\hat{B}_{\bar{P}}$ of $B_{\bar{P}}$ is isomorphic to $R[[x,y]]/(x^p y^q)$. Therefore the map $f\colon U \to \operatorname{Spec} R[x,y]/(x^p y^q)$ induces an isomorphism on the completion at $\bar{P}$, hence is étale near $\bar{P}$. Replacing U by a smaller affine, we may suppose that f is étale. Consider the diagram

$$\begin{tikzcd}
U \arrow[d, "f"'] & C \cap U \arrow[l] \arrow[d, "f_{\mathrm{red}}"] \\
\operatorname{Spec} R[x,y]/(x^p y^q) & \operatorname{Spec} R[x,y]/(xy), \arrow[l]
\end{tikzcd}$$

where the vertical arrows are étale. Since C is given, the map f_{red} is determined uniquely by the choice of x and y. Hence [7, I. 8. 3] f is also determined by these choices, and so $U = \operatorname{Spec} B$ is determined up to canonical isomorphism by elements x and y above. For the same reason, the entire formal scheme $\mathfrak{X}$ is determined up to canonical isomorphism in a neighborhood of $\bar{P}$ by the choice of elements x, y defining C^1 and C^2.

The case that $\bar{P}$ lies only on one curve $\bar{C}^1$ is similar: Let $U = \operatorname{Spec} B$ be an affine neighborhood of $\bar{P}$ such that there exist elements x, y of B, where x induces a local parameter of $\bar{C}^1$ at $\bar{P}$ and where $y = 0$ defines the closed subscheme $U \cap C^1$ of U. Then the map $f\colon U \rightarrow \operatorname{Spec} R[x, y]/(y^q)$ $(q = n_1)$ is étale near $\bar{P}$, we may suppose it is étale, and since $f_{\mathrm{red}}\colon C \cap U \rightarrow \operatorname{Spec} R[x]$ is determined by the choice of x it follows from [7, I. 5. 5] that f is determined by the choice of x and y, and hence that U is, too. Again, this reasoning also applies to the entire formal scheme $\mathfrak{X}$ near $\bar{P}$.

Now in order to calculate variations of structure by an H^1, we have first to construct some formal embedding of C into a regular scheme which will be isomorphic to $\mathfrak{X}$ locally. We could use Giraud's theory of non-abelian H^2 [4] to construct such a reference embedding, but we prefer to construct it directly as follows: Since each curve C^i is projective, so is C. Let C be embedded in projective space $\boldsymbol{P}^N{}_R$. A generic projection of $\boldsymbol{P}^N{}_k$ to $\boldsymbol{P}^2{}_k$ will induce an immersion generically on $\bar{C}$, and it is easily seen that it can be chosen so as to induce an immersion of $\bar{C}$ at a given finite set of points. Therefore we can choose *two* projections of $\bar{C}$ to $\boldsymbol{P}^2{}_k$ such that at any given point of $\bar{C}$ at least one is an immersion. The projections can be extended to projections $\boldsymbol{P}^N{}_R \rightarrow \boldsymbol{P}^2{}_R$, which induce maps $\pi_1, \pi_2\colon C \rightarrow \boldsymbol{P}^2{}_R$ which are also immersions generically. Then for every point $\bar{P}$ of the closed fibre at least one of the two maps is an immersion at $\bar{P}$, and so C can be covered by *two* open sets U_ν $(\nu = 1, 2)$ such that $\pi_\nu\colon U_\nu \rightarrow \boldsymbol{P}^2{}_R$ is an immersion of U_ν into the nonsingular space $\boldsymbol{P}^2{}_R$. The completion $\mathfrak{U}_\nu$ of $\boldsymbol{P}^2{}_R$ along U_ν is a formal scheme which is immediately seen to be locally isomorphic to what we want for $\mathfrak{X}$. Let $W = U_1 \cap U_2$. To get a formal embedding of C, we only have to find an isomorphism of $\mathfrak{U}_1 \mid W$ with $\mathfrak{U}_2 \mid W$. Now we can still make U_1 and U_2 smaller provided we keep $C = U_1 \cup U_2$.

Let $Z \subset W$ be any irreducible closed subset of dimension $\leqq 1$ not equal to a $\bar{C}^i$, and let Z' be its closure in C. Z' is finite over $\operatorname{Spec} R$, and irreducible, hence has one closed point z (by Hensel's lemma). Thus either $Z = Z'$ (in which case we can leave Z out of one of the U_ν and hence eliminate it from W), or $Z = Z' - z$. Then $Z' \cap U_\nu = Z$ for one of the two indices, say $\nu = 1$, and $Z' \cap U_\nu = Z'$ for the other index $(= 2)$. We can omit z from U_1, keeping $C = U_1 \cup U_2$. Thus in any case we can adjust U_1, U_2 to eliminate Z. It follows that W can be made a disjoint union $W = \coprod_i W^i$ of *arbitrarily small* open subsets of certain C^i which contain points of $\bar{C}^i$. Since we saw above that the formal embedding is determined locally at any point of $\bar{C}$, it follows that when W is small enough, the two formal structures $\mathfrak{U}_\nu$ will be isomorphic on each W^i, hence on W. This shows that some formal embedding of C exists.

Call the reference embedding constructed above $\mathfrak{V}$, and let V_n be the fundamental system of neighborhoods of C in defined by the ideal of function on $\mathfrak{V}$ which vanish to order at least n_i on C_i $(n = (n_1, \cdots, n_r))$. Let $\mathfrak{V}$, $\bar{V}_n$ be the reductions of $\mathfrak{V}$, V_n $(\operatorname{mod} t)$. By arguments exactly analogous to those used to show the local uniqueness of $\mathfrak{X}$, one can show that $\mathfrak{V}$ is locally isomorphic to $\mathfrak{X}$. Therefore $\mathfrak{X}$ is represented by a class $\bar{\xi}$ in $\underset{\underset{n}{\leftarrow}}{\lim} H^1(\bar{C}, \operatorname{Aut} \bar{V}_n)$, where $\operatorname{Aut} \bar{V}_n$ is the sheaf of local automorphisms of the structure $\bar{V}_n$. It is the opposite sheaf to the sheaf of groups $\operatorname{Aut}_k(\mathcal{O}_{\bar{V}_n}; \mathcal{O}_{\bar{C}})$ of local automorphisms of $\mathcal{O}_{\bar{V}}$ over k which reduce to the identity in $\mathcal{O}_{\bar{C}}$. To prove the existence of an $\mathfrak{X}$, we have to lift the class $\bar{\xi}$ to a class ξ in the set

$$\underset{\leftarrow}{\lim}\, H^1(C, \operatorname{Aut} V_n)$$

which classifies the formal schemes locally isomorphic to $\mathfrak{V}$. To do this we need the following lemma:

LEMMA (3.3). (i). *The cohomology* $H^1(C, \operatorname{Aut} V_n)$ *(resp.* $H^1(\bar{C}, \operatorname{Aut} \bar{V}_n)$*) is zero "semi-locally," i. e., an element becomes zero in a sufficiently small neighborhood of any given finite set of points of* $\bar{C}$.

(ii). *The maps* $\operatorname{Aut} V_n \to \operatorname{Aut} \bar{V}_n$ *and* $\operatorname{Aut} V_n \to \operatorname{Aut} V_{n'}$ *for* $n \geqq n'$ *are surjective.*

(iii). *For sufficiently large n, the formal scheme $\bar{\mathfrak{X}}$ is determined uniquely by the knowledge of the neighborhood $\bar{Y}_n$* (cf. Section 2).

Assume the lemma for the moment. The assertion (i) implies that any cohomology class α can be gotten as a Čech cohomology class from a covering of C by two open sets U_1, U_2 (such a class is represented by a cocycle which is just a section of $\operatorname{Aut} V_n$ (resp. $\operatorname{Aut} \bar{V}_n$) on $W = U_1 \cap U_2$). To show this, one has only to find a covering by opens U_1, U_2 such that α induces zero on U_ν. Now α is zero locally on each C^i and so there is a dense open subset U_1 of C (of the form $U_1 = \coprod_i U_1{}^i$) such that α is zero on U_1. All but a finite number of points of $\bar{C}$ lie in U_1. By (i) there is a neighborhood U_2 of this finite set on which α is zero. Since $\bar{C} \subset U_1 \cup U_2$, the open sets cover C.

Now we can replace U_ν by smaller opens so long as they still cover C. As was seen above, one can make $W = U_1 \cap U_2$ a disjoint union of arbitrarily small neighborhoods of points of some $\bar{C}^i$ in this way. Therefore using (3.1) (ii), any cohomology class s for $\operatorname{Aut} \bar{V}_n$ (resp. $\operatorname{Aut} V_{n'}$) can be lifted to $\operatorname{Aut} V_n$.

Let n be large enough so that $\bar{\mathfrak{X}}$ is determined by $\bar{Y}_n$ ((3.1) (iii)), i.e., so that the cohomology class $\bar{\xi}_n$ in $H^1(\bar{C}, \operatorname{Aut} \bar{V}_n)$ representing $\bar{V}_n$ lifts uniquely to $H^1(\bar{C}, \operatorname{Aut} \bar{V}_{n'})$ for all $n' \geqq n$. As was seen above, ξ_n can be lifted to a cohomology class $\xi_n \in H^1(C, \operatorname{Aut} V_n)$. Since also classes can be lifted from $\operatorname{Aut} V_{n'}$ to $\operatorname{Aut} V_{n''}$ for $n'' \geqq n'$, $\bar{\xi}_n$ lifts to a class of $\underset{\overset{\leftarrow}{n}}{\lim} H^1(C, \operatorname{Aut} V_n)$. The formal scheme $\mathfrak{X}$ determined by this class clearly satisfies the required conditions and so Proposition (2.2) is proved.

Proof of (3.1). Let $\bar{P}$ be a point of $\bar{C}$, say an intersection of C^1 and C^2, and revert to the notation of the beginning of this section. The sections of $\operatorname{Aut} V_n \approx \operatorname{Aut} Y_n$ on the neighborhood $U = \operatorname{Spec} B$ are the automorphisms of B over R which reduce to the identity $(\operatorname{mod} xy)$ (note that an automorphism of B necessarily carries the ideal (xy) to itself and hence induces an automorphism of $B/(xy)$). Let x', y' be elements of B with $x' \equiv x \ (\operatorname{mod} xy)$ and $y' \equiv y \ (\operatorname{mod} xy)$. Then xy, y' define the subschemes C^1, C^2 in a neighborhood of $\bar{P}$, and thus the induced map $f' \colon U \to \operatorname{Spec} R[x', y']/x'^p y'^{q'}$ is étale at $\bar{P}$ for the same reason that f was, and $f'_{\mathrm{red}} = f_{\mathrm{red}}$ (if we identify $R[x, y]$ with $R[x', y']$ by $x \rightsquigarrow x'$, $y \rightsquigarrow y'$). Therefore by [7, I. 5. 5] there is a *unique* automorphism ϕ of a neighborhood of $\bar{P}$ over $\operatorname{Spec} R$ which sends x to x' and y to y', and ϕ reduces to the identity $(\operatorname{mod} xy)$. Thus we have proved assertion (a) of the following lemma. The other assertions are analogous, and we leave them to the reader.

LEMMA (3.2). (a). *Suppose $\bar{P}$ is an intersection point of $\bar{C}^1$, $\bar{C}^2$ and let $\mathfrak{o}_n$ be the local ring of V_n at $\bar{P}$. Let $x, y \in \mathfrak{o}_n$ be local equations for C^1, C^2, at $\bar{P}$. An element ϕ of $(\operatorname{Aut} V_n)_{\bar{P}}$ is uniquely determined by the images $x' = \phi(x)$, $y' = \phi(y)$ and any elements $x' \equiv x \ (\operatorname{mod} xy)$, $y' \equiv y \ (\operatorname{mod} xy)$ are images under some ϕ.*

(ā). *Same assertions as* (a), *but with bars on all letters.*

(b). *Suppose $\bar{P}$ lies on only one curve $\bar{C}^1$. Let $\mathfrak{o}_n$ be the local ring of V_n at $\bar{P}$, and x an element of $\mathfrak{o}_n$ inducing a local parameter of $\bar{C}^1$ at $\bar{P}$, and y an element defining C^1 locally at $\bar{P}$. An automorphism $\phi \in (\operatorname{Aut} V_n)_P$ is uniquely determined by the images $x' = \phi(x)$, $y' = \phi(y)$ and any elements $x' \equiv x \ (\operatorname{mod} y)$, $y' = uy$ (u a unit) are images under some ϕ.*

($\bar{\mathrm{b}}$). *Same assertion as* (b) *but with bars on all letters.*

The surjectivity (3.1) (ii) follows immediately from this lemma, since

$\mathfrak{o}_n \to \bar{\mathfrak{o}}_n$ (resp. $\mathfrak{o}_n \to \mathfrak{o}_{n'}$) is surjective, and so once can lift the image elements of x, y under ϕ. For (3.1)(i), it suffices to show that there is a (normal) filtration of $\operatorname{Aut} V_n$ (resp. $\operatorname{Aut} \bar{V}_n$) such that the cohomology H^1 of each successive quotient is zero semi-locally. Lemma (3.2) allows us to calculate the successive quotients for a suitable filtration explicitly as follows. The calculation is exactly analogous to that done in [5] and [10]. Suppose n' is obtained from $n = (n_1, \cdots, n_r)$ by increasing one of the n_i, say n_1, by 1, and let n'' be obtained by decreasing the same index n_1 by 1 (we allow the index to become zero). It suffices to calculate the kernel $\operatorname{Aut}(V_{n'}; V_n)$ of the surjective map $\operatorname{Aut} V_{n'} \to \operatorname{Aut} V_n$, and to calculate the analogous kernel with bars. This kernel is the sheaf of local automorphsims of $V_{n'}$ which reduce to the identity on V_n. Let I be the ideal of C^1 in $V_{n'}$, which is naturally an $\mathcal{O}_{V_n}$-module. An automorphism of $V_{n'}$ induces an automorphism of I, and because of (3.2), one sees immediately that the automorphisms induced by elements of $\operatorname{Aut}(V_{n'}; V_n)$ come from units reducing to 1 in $V_{n''}$, i. e., from elements of $K = \ker(\mathcal{O}_{V_n}{}^* \to \mathcal{O}_{V_{n''}}{}^*)$. The automorphisms which leave I fixed are seen as in [5] to be isomorphic to $J \otimes T$ where $J = \ker(\mathcal{O}_{V_{n'}} \to \mathcal{O}_{V_n})$ and where T is the tangent sheaf to C^1. Thus we have an exact sequence

$$0 \to J \otimes T \to \operatorname{Aut}(V_{n'}; V_n) \to K \to 0.$$

Since the extreme terms have trivial cohomology semi-locally, so does $\operatorname{Aut} V_n$. The same calculation holds for $\operatorname{Aut} \bar{V}_n$.

There remains (3.1)(iii). This is just the rigidity statement of [5], [10], and it holds for the following reason. Since $\|(\bar{C}^i \cdot \bar{C}^j)\|$ is negative definite [11] a large enough n has the property that for at least one i, the intersection number $(\bar{C}^i \cdot \bar{Y}_n)$ will be less than a previously given negative number N. Then if we replace in the above calculation the V's by $\bar{Y}$'s, we get an exact sequence

$$0 \to \bar{J} \otimes \bar{T} \to \operatorname{Aut}(\bar{Y}_{n'}; \bar{Y}_n) \to \bar{K} \to 0$$

where

$$\bar{K} = \ker(\mathcal{O}_{\bar{Y}_n}{}^* \to \mathcal{O}_{\bar{Y}_n}{}^*) \approx \ker(\mathcal{O}_{\bar{Y}_{n'}} \to \mathcal{O}_{\bar{Y}_n}) \quad [2, \text{p. } 487],$$

$\bar{J} = \ker(\mathcal{O}_{Y_{n'}} \to \mathcal{O}_{Y_n})$, $\bar{T} =$ tangent sheaf to $\bar{C}^i$. A simple calculation shows that $\deg(\bar{J}) = -(\bar{C}^i \cdot \bar{Y}_n)$ and $\deg(\bar{K}) = -(\bar{C}^i \cdot \bar{Y}_{n''})$ [2, (1.8)]. Hence $\bar{J} \otimes \bar{T}$ and $\bar{K}$ have large degree, and so if N was chosen negative enough, $H^1(\bar{C}, \operatorname{Aut}(Y_{n'}; \bar{Y}_n)) = 0$, whence $H^1(\bar{C}, \operatorname{Aut} \bar{Y}_{n'}) = H^1(\bar{C}, \operatorname{Aut} \bar{Y}_n)$. This implies that the structure of $\bar{Y}_{n'}$ is uniquely determined by that or $\bar{Y}_n$. We can replace n by n' and repeat this process, and it is immediately seen that therefore the entire formal scheme $\bar{\mathfrak{X}}$ is determined by $\bar{Y}_n$.

4. Lifting of coverings. In this section we revert to the notation of Section 2. Suppose we are given a ring A and a map $f\colon X \to \operatorname{Spec} A$ having the properties of Theorem (2.1). We are going to show that extension of $\bar{A}$ which are tamely ramified only along the curves $\bar{C}^i$ "extend to A."

LEMMA (4.1). *Let r_i $(i = 1, \cdots, r)$ be positive integers prime to the residue characteristic p of A. There exists a finite normal Galois extension A' of A such that the normalization $f'\colon X' \to \operatorname{Spec} A'$ fo f in the function field of A' has one irreducible scheme $C^{i\prime}$ lying over each C^i, such that the ramification index of the extension along C^i is r_i, and such that the conditions* (a, b, d) *of* (2.1) *hold for $f'\colon X' \to \operatorname{Spec} A'$, with suitable intersections P'^j.*

Proof. Let D^i be the Cartier divisor on X such that $\mathcal{L}(D^i)$ and $\mathcal{L}(D^i + C^i)$ are very ample on X relative to $\operatorname{Spec} A$. For a given projective embedding of X over $\operatorname{Spec} A$, a generic linear section will intersect $\bar{C}$ transversally, and can be chosen so as to avoid any finite set of points. Hence there exist positive divisors $D^i{}_1$, $D^i{}_2$ linearly equivalent to D^i, and $H^i{}_1$, $H^i{}_2$ linearly equivalent to $D^i + C^i$ $(i = 1, \cdots, r)$ which intersect $\bar{C}$ transversally and which have no points on $\bar{C}$, (hence no points at all) in common. Now these divisors are represented by closed subschemes of X (we denote them by the same letter) which, since they intersect $\bar{C}$ transversally, are nonsingular at every point above the closed point of $\operatorname{Spec} A$, hence are nonsingular, and it is clear that their intersections with C are transversal. Let $g^i{}_\nu$ be a function whose divisor is

$$(g^i{}_\nu) = H^i{}_\nu - C^i - D^i{}_\nu$$

and let A' be the extension obtained by adjoining r_i-th roots of $g^i{}_\nu$ to A for all $\nu = 1, 2$ and $i = 1, \cdots, r$. A' is flat over A and is immediately seen to be normal by Serre's criterion. The extension A'/A is obviously Galois. A local analysis shows that the normalization X' of X in this extension is nonsingular, that there is one irreducible subscheme C'^i of X lying above C^i, and that the ramification index along C^i is r_i. Knowing this, it is easily checked that (a, b, d) of (2.1) hold for f'.

THEOREM (4.2). *With the notation of* (2.1), *let* $U = \operatorname{Spec} A - S$, $\bar{U} = \operatorname{Spec} \bar{A} - \operatorname{Spec} k = \operatorname{Spec} \bar{A}' - \operatorname{Spec} k$.

(i). *The restriction functor is an equivalence of categories between the category of étale coverings of U which are tamely ramified along each C^i, and the category of étale coverings of $\bar{U}$ tamely ramified along each $\bar{C}^i$. If*

R is an unequal characteristic ring, no étale extension of $\bar{U}$ which is wildly ramified along some $\bar{C}^i$ extends to U.

(ii). *Let K be the field of fractions of R and $\bar{K}$ the separable algebraic closure of K. Let B be the completion of the ring $A \otimes_K \bar{K}$ at the* (*unique*) *point $\bar{s}$ lying over the generic point s of S. Let U_B be the complement of the closed point of* Spec B. *There is an equivalence of categories between the categories of étale Galois coverings prime to p* (*the residue characteristic of A*) *of $\bar{U}$ and of U_B.*

Proof. (i). Since A is normal and of dimension $\geqq 3$, it follows from the local Lefschetz theorems [8, Exp. 10] that the restriction functor on étale coverings from U to $\bar{U}$ is fully faithful. Suppose an étale covering $\bar{V}$ of $\bar{U}$ tamely ramified along the $\bar{C}^i$ is given. We have to extend it to U.

We may assume that $\bar{V}/\bar{U}$ is a Galois extension. Let r_i be the ramification index of the extension along $\bar{C}^i$ and choose an extension A'/A as in (4.1). By Abyankhar's Lemma [7, X. 3. 6], the étale extension $\bar{V}'/\bar{U}'$ ($\bar{U}' = \operatorname{Spec} A'/tA' - \operatorname{Spec} K$) is unramified along any of the curves $\bar{C}'^i$, and therefore by the purity theorem [7, X. 3. 1] is induced by an étale $\bar{Y}'$ of $\bar{X}'$.

Since $\bar{X}'$, X' are proper over the complete local ring A' and have the same closed fibre, the categories of étale coverings of the two schemes are equivalent by the theorem of specialization of the fundamental group [9, 182-24, Thm. 12]. Therefore $\bar{Y}'/\bar{X}'$ is induced by an étale covering V' of U'. By descent theory [7, VIII], it suffices to find descent data for the covering $\bar{V}'/\bar{U}'$ relative to the map $U \leftarrow U'$. We have descent data for the covering $\bar{V}'/\bar{U}$ relative to the map $\bar{U} \leftarrow \bar{U}'$. Therefore it suffices to show that with the obvious notation the restriction functor from U'' to $\bar{U}''$ (resp. U''' to $\bar{U}'''$) is fully faithful (resp. faithful), where $A'' = A' \otimes_A A'$ and $A''' = A' \otimes_A A' \otimes_A A'$. Thus we are led by [8, X, (2.1), (2.3)] to verify that the depth condition of [8, X. 2. 1(2)] holds for A'' and A'''. Since A is normal and of dimension 3, it holds for A. Since A'', A''' are flat over A, it holds for them as well. This shows that $\bar{V}/\bar{U}$ is induced from an étale extension V of U.

To complete the proof of (i), it remains to show that no étale covering wildly ramified along a curve $\bar{C}^i$ extends, if R is an unequal characteristic ring. Let V/U be an étale extension, and let $\bar{P}$ be a simple point of $\bar{C}$. Let B be the completion of the local ring of X at $\bar{P}$ and let B'/B be the normal extensions induced by V/U. B'/B is ramified only along the locus $\{x = 0\}$ where $x = 0$ is a defining equation for C at P, which is a local parameter in B.

We want to show that the ramification index of B'/B along $x=0$ is prime to p. The ring B' is a product of local rings and we may replace it by one of its factors. Thus we are reduced to the following lemma:

LEMMA (4.3). *Let B be a complete regular unequal characteristic local ring with separably closed residue field of characteristic p, and let x be a local parameter in B which does not divide p. Let B'/B be an irreducible normal extension ramified only along the locus $\{x=0\}$. Then* $\deg(B/A)$ *is prime to p.*

Proof. We may assume the extension Galois. Let n be the ramification index of the extension at the generic point of $x=0$. Since $x \nmid p$, this point has residue characteristic zero. Put $C=B[\xi]$ where ξ satisfies the equation $\xi^n=x$. Then C is obviously a regular local ring. Let C' be the normalization of $B'\otimes_B C$. By Abyankhar's lemma, C'/C is unramified at the generic points of $\{\xi=0\}$ in $\operatorname{Spec} C'$, therefore unramified in codimension 1, whence unramified by purity. Since the residue field of C is sepabably closed, C'/C is the trivial (completely decomposed) extension. Therefore since $n \leqq \deg(B'/B)$, the extensions C and B' of B are isomorphic. But if n were not prime to p, the extension C/B would be ramified along the locus $\{p=0\}$ in $\operatorname{Spec} B$, which is not the case by assumption. Therefore $(n,p)=1$.

It remains to prove (4.2)(ii). Consider the base change functor from étale coverings of U to étale coverings of U_B. If we show that it is an equivalence of categories on Galois coverings prime to p, (ii) will follow, using (i). Now to show that this functor is fully faithful, it suffices by the usual argument of the graph to show that if V/U is a connected étale covering, then the induced covering V_B/U_B is also connected. Now U_B is irreducible and regular. For, it is easily seen that $X_B=X\otimes_A B$ is regular and connected, hence irreducible, and U_B is isomorphic to an open subset of X_B. Therefore to show V_B irreducible is the same as to show it integral. Let A'/A be an extension as in (4.1) such that the induced covering V'/U' is unramified along each of the C^i. Let V'' be an irreducible component of V'. Then V'' is a (ramified) Galois extension of U having V as a subextension and so it suffices to show that V''_B is irreducible. Now $U'_B=U_B\otimes_A A'$ is the complement of the closed point of $\operatorname{Spec} B'$, where $B'=B\otimes_A A'$ is the completion of $A'\otimes_K \bar{K}$ at the point above s, and hence U'_B is irreducible for the same reason that U_B is. Therefore we may as well replace A by A' and V by V''. This means that we are reduced to the case that the extension V/U is unramified along any of the C^i, i. e., is the restriction of a connected étale covering of X to U. Since this covering is connected, its reduced closed fibre, which is an étale covering Y of C is. We need only show that the induced covering Y_B of C_B is also connected, which follows immediately from the fact that C is the union of smooth curves over $\operatorname{Spec} R$ with normal crossings and that C_B is a geometric fibre of $C/\operatorname{Spec} R$.

It remains to show that every étale Galois covering V_B of U_B prime to p is induced by covering V/U. Let again A'/A be as in (4.1), and let V'_B/U'_B be the induced extension. If we show that an irreducible component V''_B of V'_B is induced by an étale covering V'' of U', then V''/U will be a (ramified) Galois extension by the above argument, and hence the subextension V_B/U_B of V''_B will be induced from a subextension V of V''/U. Moreover it is immediately seen that V/U must be étale since V_B/U_B is. Thus we may replace A by the extension A'. In this way we can reduce ourselves to the case that V_B/U_B is unramified along any of the curves $C^i{}_B$, i.e., comes from an étale covering of X_B. Let Y_B/C_B be the induced covering. By the theorem of specialization of the fundamental group [9, 182-26, Thm. 13] (slightly generalized to include the case of a union of smooth curves with crossings) this extension is induced by an étale covering Y of C, which is induced [9, 182-24, Thm. 12] by an étale covering of X, yielding the desired extension V/U. This completes the proof of (4.2).

5. Mumford's example. Theorem (4.2) (ii) will allow one to calculate the quotient "prime to p" of the fundamental group of $\bar{U}$ by the topological method of [12], once one has sufficient facility with change of ground field and passage to the completion. However according to (4.1) (i) no wildly ramified extension can be lifted. In view of this it would be very interesting to obtain some finiteness statement for the fundamental groups including the p-extensions.

We want to close with an example of *Mumford* which shows that extensions wildly ramified along the C^i do indeed occur:

Consider the singularity at the origin in $\operatorname{Spec} k[[x,y,z]]$ where x, y, z satisfy the equation

$$z^2 + a\,xyz + x^2y + xy^2 = 0 \tag{5.1}$$

for some $a \in k$. If the characteristic of k is not 2, the fundamental group of this local scheme (minus the closed point), U, is the quaternion group. If the characteristic is 2, however, the fundamental group is $\mathbf{Z}/2$ when $a \neq 0$, and is trivial when $a = 0$. The nontrivial extension is given by $k[[u,v]]$ where

$$u^2 + axu + x = 0$$
$$v^2 + ayv + y = 0$$
$$uy + vx = z.$$

If $a \neq 0$, this extension is étale except at the closed point, and it is easily seen that the extension is ramified along all prime divisors above this point, and therefore does not lift to characteristic zero. If $a = 0$, the extension is purely inseparable, and hence induces an equivalence of categories between thet étale extensions of the two schemes. Since $\operatorname{Spec} k[[u, v]]$ is regular, this shows that the fundamental group of U is trivial in that case.

The singularities (5.1) are interesting from other points of view. If the characteristic is not 2 ,the singuarity is rigid, i. e., does not depend on a (cf. Kirby [11]), while in characteristic 2 they are not all equivalent, as was seen above. Equation (5.1) actually describes an "equisingular" deformation over $\operatorname{Spec} Z$, say for $a = 1$. It is instructive to analyze the behavior of its branch locus over $\operatorname{Spec} \mathbf{Z}[x, y]$ at the prime 2.

REFERENCES.

[1] S. Abyankhar, "Local uniformization on algebraic surfaces over ground fields of characteristic $p \neq 0$," *Annals of Mathematics*, vol. 63 (1956), pp. 491-526.

[2] M. Artin, "Some numerical criteria for contractability of curves on algebraic surfaces," *American Journal of Mathematics*, vol. 84 (1962), pp. 485-496.

[3] N. Bourbaki, *Éléments de Mathématique*, Algèbre Commutatif, Chapter 3, Paris (1961).

[4] J. Giraud, Thesis, Paris (forthcoming).

[5] H. Grauert, "Über Modifikationen und exzeptionelle analytische Mengen," *Mathematische Annalen*, vol. 146 (1962), pp. 369-374.

[6] A. Grothendieck and J. Dieudonné, *Éléments de Géométrie Algébrique*, Institut des Hautes Études Scieitifiques, Pub. Math., No. 4 (1960-).

[7] A. Grothendieck, *Séminaire de Géométrie Algébrique*, Institut des Hautes Études Scientifiques, Mimeographed Notes (1960-61).

[8] ———, *Séminaire de Géométrie Algébrique*, Institut des Hautes Études Scientifiques, Mimeographed Notes (1962).

[9] ———, *Fondements de Géométrie Algébrique*, Mimeographed Notes, Paris (1962),

[10] H. Hironaka and H. Rossi, "On the equivalence of embeddings of exceptional complex spaces," *Mathematische Annalen*, vol. 156 (1964), pp. 313-333.

[11] D. Kirby, "The structure of an isolated multiple point of a surface, II," *Proceedings of the London Mathematical Society*, series 3, vol. 7 (1957), pp. 118.

[12] D. Mumford, *The topology of normal singularities of an algebraic surface and a criterion for simplicity*, Institut des Hautes Études Scientifiques, Pub. Math. 9 (1960).

[13] O. Zariski, "The reduction of singularities of an algebraic surface," *Annals of Mathematics*, vol. 40 (1939), pp. 639-689.

ON THE STRUCTURE OF LOCAL ORTHOGONAL GROUPS.

By Barth Pollak.*

The structure of the projective commutator subgroup of the orthogonal group is well known when isotropic lines are present. The situation is radically different in the anisotropic case where results are meager. In [3] Dieudonné gives an example of a 3-dimensional anisotropic quadratic space over a p-adic field in which the commutator subgroup Ω of the orthogonal group contains an infinite descending chain

$$\Omega = \Gamma_0 \supset \Gamma_1 \supset \Gamma_2 \supset \cdots \supset \Gamma_n \supset \cdots$$

of normal subgroups (the "congruence groups") such that Γ_n/Γ_{n+1} is an abelian p-group and $\bigcap_{n=1}^{\infty} \Gamma_n = 1$. (See also [1, pp. 179-186], [4], [5, p. 57], [6]). In a recent visit to Notre Dame, Dieudonné raised with me the problem of determining all the normal subgroups of Ω. It is the purpose of the present paper to make this determination. A brief description now follows.

We refine Dieudonné's chain by interpolating, for each non-negative integer n, a certain normal subgroup Δ_n of Ω between Γ_{n+1} and Γ_n, i. e. $\Gamma_{n+1} \subset \Delta_n \subset \Gamma_n$. We then show that, if Γ is an arbitrary non-trivial normal subgroup of Ω, there exists a non-negative integer m such that

$$\Delta_m \subseteq \Gamma \subseteq \Gamma_m.$$

Moreover, any subgroup Γ of Ω satisfying $\Delta_m \subseteq \Gamma \subseteq \Gamma_m$ is normal in Ω. Thus all normal subgroup of Ω lie in a top layer of the refined chain. The bottom layers $\Gamma_{n+1} \subset \Delta_n$ (excluding Γ_{n+1} and Δ_n) are devoid of normal subgroups of Ω. There is one exception to this rule. It occurs when $p = 3$. Then the roles of the top and bottom layers are reversed. Thus the normal subgroups of Ω lie in the bottom layers while the top layers (excluding the "ends") are devoid of normal subgroups of Ω. (Actually we prove this not only for p-adic fields but, more generally, for any nondyadic local field).

We now give a precise description introducing the notation that we shall use in the sequel.

Received July 30, 1965.

* This work has been supported in part by the National Science Foundation under grant GP-3986.

Let V be a 3-dimensional anisotropic quadratic space over a nondyadic local field F. Let $\mathfrak{o}$ denote the valuation ring of F, $\mathfrak{p}$ the maximal ideal of $\mathfrak{o}$ and $\mathfrak{F}$ the residue class field of F. Let $O(V)$ denote the orthogonal group of V and $\Omega(V)$ the commutator subgroup of $O(V)$. There is a base for V such that, for each $\sigma \in O(V)$ the matrix of σ with respect to this base, denoted mat σ, has all of its entries in $\mathfrak{o}$. Set $\Lambda_0(V) = \Omega(V)$ and for each positive integer n set $\Lambda_{2n}(V) = \Gamma_n(V)$. Thus

$$\Lambda_{2n}(V) = \{\sigma \in O(V) \mid \operatorname{mat}\sigma \equiv E (\operatorname{mod} \mathfrak{p}^n)\}$$

where E denotes the identity matrix. Then it turns out that $\Lambda_{2n}(V) \subset \Omega(V)$ and

$$\sigma \in \Lambda_{2n}(V) \text{ implies } \operatorname{Tr}\sigma \equiv 3 \ (\operatorname{mod} \mathfrak{p}^{2n}).$$

Set

$$\Lambda_1(V) = \{\sigma \in O(V) \mid \operatorname{Tr}\sigma \equiv 3 \ (\operatorname{mod} \mathfrak{p})\}$$

and

$$\Lambda_{2n+1}(V) = \{\sigma \in \Lambda_{2n}(V) \mid \operatorname{Tr}\sigma \equiv 3 \ (\operatorname{mod} \mathfrak{p}^{2n+1})\}.$$

Again $\Lambda_{2n-1}(V) \subset \Omega(V)$ and we obtain an infinite descending chain of subgroups

$$\Omega(V) = \Lambda_0(V) \supset \Lambda_1(V) \supset \cdots \supset \Lambda_n \supset \cdots$$

and $\bigcap_{n=0}^{\infty} \Lambda_n(V) = 1$.

Our main result can now be stated.

THEOREM. *For each non-negative integer n, $\Lambda_n(V)$ is a normal subgroup of $\Omega(V)$. For the factor groups we have first of all*

$$\Lambda_0(V)/\Lambda_1(V) \cong \mathfrak{N}/\{-1,1\}$$

where $\mathfrak{N}$ is the group of elements of norm 1 in the (unique) quadratic extension of $\mathfrak{F}$. If $n > 0$ then $\Lambda_{2n}(V)/\Lambda_{2n+1}(V) \cong \mathfrak{F}^+$, the additive group of $\mathfrak{F}$,

$$\Lambda_{2n-1}(V)/\Lambda_{2n}(V) \cong \mathfrak{F}^+ \oplus \mathfrak{F}^+.$$

1) *If $\mathfrak{F} \neq \boldsymbol{F}_3$, the field of 3 elements, then every non-trivial normal subgroup Γ of $\Omega(V)$ satisfies*

$$(*) \qquad \Lambda_{2n+1}(V) \subseteq \Gamma \subseteq \Lambda_{2n}(V)$$

for some n. Moreover, every subgroup of $\Omega(V)$ satisfying (*) is normal in $\Omega(V)$.

2) *If $\mathfrak{F} = \boldsymbol{F}_3$ then every non-trivial normal subgroup Γ of $\Omega(V)$ satisfies*

$$(^{**}) \qquad \Lambda_{2n}(V) \subseteq \Gamma \subseteq \Lambda_{2n-1}(V)$$

for some n. Moreover, every subgroup of $\Omega(V)$ *satisfying* (**) (*there are* 6 *such*) *is normal in* $\Omega(V)$.

Note that in the case $\mathfrak{J} = \boldsymbol{F}_p$, p a prime greater than 3, the only normal subgroups of $\Omega(V)$ are the $\Lambda_n(V)$.

We shall use the spinorial representation of $O^+(V)$ to prove this result. The main tools are several theorems of the author dealing with quaternion algebras (see § 5). The proof itself is given in §§ 6, 7 and 8. The crucial step is to show that every non-trivial normal subgroup Γ is "sandwiched" between two successive Λ's.

Our result further elucidates Jacobson's example concerning the multiplicative group of a division ring (see [7], p. 191) although it appears that the choice $p = 3$ in his example results in exceptional behavior.

1. Three-dimensional quadratic spaces.

In this section we wish to set forth in our language and notation various well known facts concerning the orthogonal group of a 3-dimensional quadratic space and in particular its spinorial representation.

Let F be a (commutative) field of characteristic $\neq 2$ and A a quaternion algebra over F. Then A is 4-dimensional central simple over F and has a base of the form $1, i, j, ij$ with $i^2 = \alpha$, $j^2 = \beta$, $ij = -ji$ where $\alpha, \beta \in F^*$ (the multiplicative group of nonzero elements of F). We shall use the notation (α, β) for a quaternion algebra possessing such a base. A possesses an anti-automorphism of period 2 called *conjugation*, the image of $t \in A$ being denoted $\bar{t}$. We have $t + \bar{t} = S(t)1$, $t\bar{t} = N(t)1$ with $S(t)$, $N(t) \in F$ called respectively the *trace* and *norm* of t and $t^2 - S(t)t + N(t)1 = 0$ for each $t \in A$. If $S(t) = 0$ we call t *pure*. If $t = t_0 1 + t_1 i + t_2 j + t_3 ij$ we have

$$N(t) = t_0^2 - \alpha t_1^2 - \beta t_2^2 + \alpha\beta t_3^2. \tag{1}$$

Now let V be a 3-dimensional vector space over F upon which is defined a nonsingular quadratic form f. Let $O(V)$ denote the orthogonal group of V, $O^+(V)$ the subgroup of elements of determinant $+1$ (called *rotations*), $O'(V)$ the spinorial kernel and $\Omega(V)$ the commutator subgroup of $O(V)$. Since $\dim V = 3$, $O'(V) = \Omega(V)$ (see [10], 55:5).

We denote by τ_a the symmetry with respect to the hyperplane perpendicular to the anisotropic vector a and remark that every $\sigma \in O^+(V)$ is of the form $\sigma = \tau_a \tau_b$ (see [10], 43:3). We note that replacing f by γf, $\gamma \in F^*$, leaves $O(V)$ unchanged and hence if we choose a base for V and write

$f=\alpha_1X_1^2+\alpha_2X_2^2+\alpha_3X_3^2$, replace f by $(\alpha_1\alpha_2\alpha_3)f$ and set $-\alpha=\alpha_2\alpha_3$, $-\beta=\alpha_3\alpha_1$, we may assume f has the form

$$f=-\alpha t_1^2-\beta t_2^2+\alpha\beta t_3^2. \tag{2}$$

Comparing this with (1), we see that there is no loss of generality in assuming that V is the set of pure quaternions in the algebra A. We also note that $\tau_a(x)=-axa^{-1}$ and hence for each $\sigma\in O^+(V)$ we have $\sigma(x)=txt^{-1}$ where $t=ab$ if $\sigma=\tau_a\tau_b$. We can now easily derive

1.1. *Let A^* denote the set of invertible quaternions of A and*

$$G=\{t\in A\mid N(t)=1\}.$$

Then there exists an exact sequence

$$1\to F^*\to A^*\xrightarrow{\phi}O^+(V)\to 1$$

where the epimorphism ϕ is given by $\phi(t)=\sigma$ and

$$\sigma(x)=txt^{-1}. \tag{3}$$

Furthermore, if $\theta(\sigma)$ denotes the spinor norm of σ, then

$$\theta(\sigma)=N(t)F^{*2}. \tag{4}$$

Restricting ϕ to G we obtain the exact sequence

$$1\to\{-1,1\}\to G\to\Omega(V)\to 1. \tag{5}$$

As $\phi(t)\in\operatorname{End}V$ we may consider $\operatorname{Tr}\phi(t)$. A straightforward computation using (3) immediately yields

1.2. *If $t\in A^*$ then*

$$\operatorname{Tr}\phi(t)=\frac{S(t)^2}{N(t)}-1.$$

Although it will not be used in the sequel, the following formula is very useful for computing the spinor norm of an isometry given in matricial terms.

1.2a. *Let $\sigma\in O^+(V)$ and suppose $1+\operatorname{Tr}\sigma\neq 0$. Then*

$$\theta(\sigma)=(1+\operatorname{Tr}\sigma)F^{*2}.$$

Proof. Write $\sigma=\phi(t)$ for some $t\in A^*$. Then by 1.2, $1+\operatorname{Tr}\sigma=N(t)(S(t)/N(t))^2$ and the result follows by (4). q. e. d.

2. Specialization of the field F. In this section as well as in the remainder of the paper we assume that F is a *nondyadic local field.* Thus F is complete with respect to a non-archimedean valuation with finite residue class field $\mathfrak{F}$ of characteristic $\neq 2$. Let $\mathfrak{o}$ denote the valuation ring of F and $\mathfrak{p}$ the prime ideal of $\mathfrak{o}$. We shall let π denote a fixed prime element of $\mathfrak{o}$ (thus $\mathfrak{p}=\pi\mathfrak{o}$), ϵ a fixed nonsquare unit of $\mathfrak{o}$. If $\gamma\in F^*$, $|\gamma|=|\pi|^{\nu}$ for some integer ν. Set $\nu=\operatorname{ord}\gamma$. If $\alpha,\beta\in\mathfrak{o}$ then $\alpha\sim\beta$ will mean $\alpha=\zeta^2\beta$ for some ζ in $\mathfrak{o}$. By the Local Square Theorem (see [10], 63:1), if α and β are units then $\alpha\sim\beta$ if and only if $\alpha/\beta\equiv$ square $(\operatorname{mod}\mathfrak{p})$.

Let A be a quaternion algebra over F. Then A is a division algebra if and only if $A\cong(\epsilon,\pi)$ where $\cong$ means "is isomorphic to" (see [10], 63:11b).

Set $J=\{t=t_0 1+t_1 i+t_2 j+t_3 ij\in A \mid t_0,\cdots,t_3\in\mathfrak{o}\}$. Then we have

2.1. *Let A be a division algebra, $t\in A$. Then*

i) *$t\in J$ if and only if $Nt\in\mathfrak{o}$.*

ii) *$t\in\pi J$ if and only if $Nt\in\mathfrak{p}^2$.*

Proof. Necessity is trivial. Sufficiency is a well known result in the representation theory of quadratic forms.

2.1a. COROLLARY. $G\subseteq J$.

2.2. *Let A be a division algebra and $\beta\in F$. Then $\beta=N(x)$ for some pure $x\in A$ if and only if $-\beta$ is a nonsquare in F.*

Proof. $N(x)$ is a ternary quadratic form of index 0 and discriminant 1. The result is immediate from [10], 63:21.

The following result is basic for what follows.

2.3. *Let A be a division algebra, $\gamma\in F$ with $\gamma\neq 1$. There exists $t\in G$ such that*

i) $S(t)=2\gamma$ ii) $t-1\in\pi^n J$

if and only if

$$\gamma=1+\pi^{2n+k}\alpha$$

where α is a unit, k is a non-negative integer and 2α is a nonsquare if k is even. Furthermore $t-1\notin\pi^{n+1}J$ if and only if $k=0$ or 1.

Proof. 1) Necessity. Let $t-1\in\pi^n J$. We may write

$$t=(1+\pi^n t_0)1+\pi^n x$$

where x is pure. (Note that $\gamma\neq 1$ implies $t_0\neq 0$). Thus

$$1 = Nt = (1 + \pi^n t_0)^2 + \pi^{2n} N(x).$$

Hence $\pi^n \mid t_0$ and we may write $t_0 = \pi^{n+k}\alpha$ for some non-negative integer k and some unit α. Then

$$N(x) = -(2\pi^k\alpha + \pi^{2n+2k}\alpha^2).$$

By 2.2, $2\pi^k\alpha + \pi^{2n+2k}\alpha^2$ is a nonsquare in F. Thus 2α is a nonsquare if k is even. By 2.1 we see that $t - 1 \notin \pi^{n+1}J$ if and only if $k = 0$ or 1.

2) Sufficiency. $-(2\pi^k\alpha + \pi^{2n+2k}\alpha^2) = N(x)$ for some pure $x \in A$ by 2.2. And $x \in J$ by 2.1. Set $t = (1 + \pi^{2n+k}\alpha)1 + \pi^n x$. Then $t - 1 \in \pi^n J$, $S(t) = 2\gamma$ and $t \in G$ as desired. q. e. d.

3. The congruence groups. From now on we suppose that V is anisotropic. Then we may take $\alpha = \epsilon$ and $\beta = \pi$ in (2) and $A = (\epsilon, \pi)$ is a division algebra. With respect to the base i, j, ij for V every $\sigma \in O(V)$ has a matrix representation, denoted mat σ, which possesses entries in $\mathfrak{o}$. (See [5], p. 57). For each positive integer n set

$$\Gamma_n(V) = \{\sigma \in O(V) \mid \operatorname{mat}\sigma \equiv E \pmod{\mathfrak{p}^n}\}$$

where E is the identity matrix and congruence of matrices is entry-wise. Although one can easily deduce from the definition that $\Gamma_n(V)$ is a normal subgroup of $O(V)$ we shall obtain this fact indirectly in a moment. At this point we content ourselves with the following.

3.1. *For each positive integer n,*

$$\Gamma_n(V) \subseteq \Omega(V).$$

Proof. Let $\sigma \in \Gamma_n(V)$. Then $\det(E + \operatorname{mat}\sigma) \equiv 8 \pmod{\mathfrak{p}}$ whence $2^3 \det(E + \operatorname{mat}\sigma)^{-1}$ is a square by the Local Square Theorem. By the formula of Lipschitz (see [9], p. 250), $\theta(\sigma) = 1$ whence $\sigma \in \Omega(V)$. q. e. d. (1.2a provides an alternative proof).

3.2. *Let $\sigma \in \Gamma_n(V)$. Then*

$$\operatorname{Tr}\sigma \equiv 3 \pmod{\mathfrak{p}^{2n}}.$$

Proof. By 1.1, there exists $t \in G$ such that $\phi(t) = \sigma$. By 1.2, $\operatorname{Tr}\sigma = S(t)^2 - 1$. The result now follows from 2.3. q. e. d.

Now set $\Lambda_{2n}(V) = \Gamma_n(V)$ for n a positive integer and define

$$\Lambda_{2n+1}(V) = \{\sigma \in \Gamma_n(V) \mid \operatorname{Tr}\sigma \equiv 3 \pmod{\mathfrak{p}^{2n+1}}\}.$$

Finally define

$$\Lambda_0(V) = \Omega(V)$$
$$\Lambda_1(V) = \{\sigma \in O(V) \mid \operatorname{Tr}\sigma \equiv 3 \pmod{\mathfrak{p}}\}.$$

As in 3.1 we see that $\Lambda_1(V) \subset \Omega(V)$ and we have an infinite descending sequence

$$\Omega(V) = \Lambda_0(V) \supseteq \Lambda_1(V) \supseteq \Lambda_2(V) \supseteq \cdots.$$

We note that $\bigcap_{n=0}^{\infty} \Lambda_n(V) = 1$. As $\Omega(V)$ is double covered by G through the epimorphism ϕ we wish to determine the preimages of the various $\Lambda_n(V)$.

Define $G_0 = G$ and

$$G_{2n} = \{t \in G \mid \text{either } t - 1 \text{ or } t + 1 \in \pi^n J\}.$$

If $t \in G_{2n}$ define $\chi(t) = 1$ or -1 by the condition $t - \chi(t)1 \in \pi^n J$ and note that $(t+1) - (t-1) = 2$ implies $\chi(t)$ is uniquely defined by t.

We have the following formulas.

If $t, u \in G_{2n}$ then

$$(6)\quad \begin{cases} tu - \chi(t)\chi(u)1 = (t-\chi(t)1)(u-\chi(u)1) + \chi(u)(t-\chi(t)1) \\ \qquad\qquad\qquad\qquad\qquad\qquad + \chi(t)(u-\chi(u)1). \\ \text{If } s \in G \text{ then} \\ sts^{-1} - \chi(t)1 = s(t-\chi(t)1)s^{-1} = s(t-\chi(t)1)\bar{s}. \end{cases}$$

Thus $\chi(tu) = \chi(t)\chi(u)$ and we have

3.3. *For each non-negative integer n, G_{2n} is a normal subgroup of G.*

Proof. Immediate by (6). q. e. d.

3.4. *Using the epimorphism ϕ of* 1.1 *we have the exact sequence*

$$1 \to \{-1, 1\} \to G_{2n} \xrightarrow{\phi} \Lambda_{2n}(V) \to 1.$$

Proof. The only assertion that needs proof is the surjectivity of ϕ. Clearly $\phi(G_{2n}) \subseteq \Lambda_{2n}(V)$. Using an argument of Dieudonné (see [3], p. 36) there exists a Cayley parametrization

$$\operatorname{mat}\sigma = (E - X)(E + X)^{-1}$$

where the entries of X lie in $\mathfrak{p}^n$. But then, using formula (3) of Lipschitz [9], p. 248, we obtain $t \in G_{2n}$ such that $\phi(t) = \sigma$. q. e. d.

By 2.3, for $n > 0$, $t \in G_{2n}$ implies $S(t) \equiv 2\chi(t) \pmod{\mathfrak{p}^{2n}}$. For $n > 0$ define

$$G_{2n+1} = \{t \in G_{2n} \mid S(t) \equiv 2\chi(t) \,(\mathrm{mod}\, \mathfrak{p}^{2n+1})\}.$$

For $n = 1$ define

$$G_1 = \{t \in G \mid S(t) \equiv 2 \text{ or } -2 \,(\mathrm{mod}\, \mathfrak{p})\}$$

and define $\chi(t) = 1$ or -1 for $t \in G_1$, by

$$S(t) \equiv 2\chi(t) \,(\mathrm{mod}\, \mathfrak{p}).$$

3.5. *For each non-negative integer n, G_{2n+1} is a normal subgroup of G.*

Proof. The only point in question is the assertion: $t, u \in G_{2n+1}$ implies $tu \in G_{2n+1}$. To see this write

$$t = (\chi(t) + \pi^{2n+k}\alpha)1 + \pi^n(t_1 i + t_2 j + t_3 ij)$$

where $k > 0$. Now $Nt = 1$ implies $t_1 \in \mathfrak{p}$. A similar remark holds for u. By direct computation one now finds that $tu \in G_{2n+1}$. q. e. d.

3.6. *Let n be a non-negative integer. Then, using the epimorphism ϕ of* 1.1, *we have an exact sequence*

$$1 \to \{-1, 1\} \to G_{2n+1} \xrightarrow{\phi} \Lambda_{2n+1}(V) \to 1.$$

Proof. This follows immediately from 1.2 and 3.4. q. e. d.

We have now determined the preimages of the $\Lambda_n(V)$ and, incidentally, demonstrated that the $\Lambda_n(V)$ are normal subgroups of $O(V)$.

4. More concerning the congruence groups. In this section we will determine the structure of the factor groups $\Lambda_n(V)/\Lambda_{n+1}(V)$. It will be convenient to introduce the notation ρ for the canonical homomorphism of $\mathfrak{o}$ onto $\mathfrak{F}$, the residue class field of F. Let $\mathfrak{F}^+$ denote the additive group of $\mathfrak{F}$.

4.1. *Let n be a positive integer. Then*

$$\Lambda_{2n}(V)/\Lambda_{2n+1}(V) \cong \mathfrak{F}^+.$$

Proof. By 3.4 and 3.6 we will instead prove $G_{2n}/G_{2n+1} \cong \mathfrak{F}^+$. Let $t \in G_{2n}$. In view of 2.3 we may write

$$t = (\chi(t) + \pi^{2n+k}t_0)1 + \pi^n(t_1 i + t_2 j + t_3 ij).$$

Set $g(t) = \rho(\chi(t)t_1)$. By direct computation we verify that, if $u \in G_{2n}$, then $g(tu) = \rho(\chi(t)t_1) + \rho(\chi(u)u_1)$ and g is a homomorphism of G_{2n} into $\mathfrak{F}^+$. The condition $Nt = 1$ implies $k > 0$ if and only if $t_1 \in \mathfrak{p}$. Thus $\ker g = G_{2n+1}$. Finally we prove surjectivity. Let $c \in \mathfrak{F}^+$. We may assume $c \neq 0$. Let $\rho(\gamma) = c$. The quadratic equation

$$\pi^{2n}X^2 + 2X - \epsilon\gamma^2 = 0$$

has square discriminant by the Local Square Theorem. Let $X = \alpha$ be the integral solution. Set $t = (1 + \pi^{2n}\alpha) + \pi^n\gamma i$. Then $t \in G_{2n}$ and $g(t) = \rho(\gamma) = c$. q. e. d.

4. 2. *Let n be a positive integer. Then*

$$\Lambda_{2n-1}(V)/\Lambda_{2n}(V) \cong \mathfrak{F}^+ \oplus \mathfrak{F}^+.$$

Proof. We shall prove that $G_{2n-1}/G_{2n} \cong \mathfrak{F}^+ \oplus \mathfrak{F}^+$. Let $t \in G_{2n-1}$. In view of 2. 3 we may write

$$t = (1 + \pi^{2n-2+k}t_0)1 + \pi^{n-1}(t_1 i + t_2 j + t_3 ij)$$

and $k \geqq 1$. Note that $Nt = 1$ implies $t_1 \in \mathfrak{p}$. Set $h(t) = \chi(t)(\rho(t_2), \rho(t_3))$. By direct computation we verify that h is a homomorphism of G_{2n-1} into $\mathfrak{F}^+ \oplus \mathfrak{F}^+$. Clearly $G_{2n} \subseteq \ker h$. If $h(t) = (0, 0)$ then, since $t_1 \in \mathfrak{p}$, $Nt = 1$ implies $k \geqq 2$ whence $t \in G_{2n}$ and $\ker h = G_{2n}$. Finally we prove that h is surjective. Let $(c_2, c_3) \in \mathfrak{F}^+ \oplus \mathfrak{F}^+$. Let $\rho(\gamma_2) = c_2$, $\rho(\gamma_3) = c_3$ and set

$$N(\gamma_2 j + \gamma_3 ij) = \pi E.$$

The quadratic equation

$$\pi^{2n}X^2 + 2\pi X + \pi E = 0$$

has square discriminant by the Local Square Theorem. Let $X = \alpha$ be the integral solution. Set.

$$t = (1 + \pi^{2n-1}\alpha)1 + \pi^{n-1}(\gamma_2 j + \gamma_3 ij).$$

Then $t \in G_{2n-1}$ and $h(t) = (\rho(\gamma_2), \rho(\gamma_3)) = (c_2, c_3)$. q. e. d.

We are left with only the group $\Lambda_0(V)/\Lambda_1(V)$ to compute. To do this first note that $\rho(\epsilon)$ is a nonsquare in $\mathfrak{F}$ by the Local Square Theorem. Hence $\mathcal{E} = \mathfrak{F}(\rho(\epsilon)^{\frac{1}{2}})$ is a quadratic extension of $\mathfrak{F}$. Set $\mathfrak{N} = \{\gamma \in \mathcal{E} \mid N_{\mathcal{E}/\mathfrak{F}}(\gamma) = 1\}$.

4. 3. $\Lambda_0(V)/\Lambda_1(V) \cong \mathfrak{N}/\{-1, 1\}$.

Proof. We have $G_0/G_1 \cong \Lambda_0(V)/\Lambda_1(V)$. Let $\mu \in \mathcal{E}$ satisfy $\mu^2 = \rho(\epsilon)$. If $t \in G_0$, say $t = t_0 1 + t_1 i + t_2 j + t_3 ij$, set $g(t) = \rho(t_0) + \rho(t_1)\mu$. Since $N(t) = 1$, $g(t) \in \mathfrak{N}$. It is straightforward to verify that g is a homomorphism. g is surjective. To see this, let $a_0 + a_1\mu \in \mathfrak{N}$. Then $a_0 + a_1\mu$ can be refined to give an element $\gamma_0 + \gamma_1 i \in G$ such that $\rho(\gamma_0) = a_0$, $\rho(\gamma_1) = a_1$. (See [2], Theorem 2 on p. 131 or alternatively, [8], Theorem 9a). Thus $g(\gamma_0 + \gamma_1 i) = a_0 + a_1\mu$. Finally we compute $\ker g$. It is easy to see that

$\ker g = \{t \in G_1 \mid \chi(t) = 1\}$. Thus $G_0/\ker g \cong \mathfrak{N}$. As $G_1/\ker g \cong \{-1, 1\}$ (χ is a homomorphism by the equation immediately following (6)) the result immediately follows. q. e. d.

We summarize the preceding results as follows.

4.6. THEOREM. *Let V be a 3-dimensional anisotropic quadratic space over a nondyadic local field F. Let $\mathfrak{F}$ be the residue class field of F, $\mathfrak{F}^+$ its additive group and $\mathfrak{N}$ the elements of the (unique) quadratic extension of $\mathfrak{F}$ of norm 1. Then there exists a sequence $\{\Lambda_n(V)\}$ $(n = 0, 1, \cdots)$ of normal subgroups of $\Omega(V)$ which satisfy the following conditions:*

a) $\Omega(V) = \Lambda_0(V)$.

b) *For each positive integer n,* $\Lambda_n(V) \subset \Lambda_{n-1}(V)$.

c) $\bigcap_{n=0}^{\infty} \Lambda_n(V) = 1$.

d) $\Lambda_0(V)/\Lambda_1(V) \cong \mathfrak{N}/\{-1, 1\}$.

e) *For each positive integer n,*
$\Lambda_{2n}(V)/\Lambda_{2n+1}(V) \cong \mathfrak{F}^+$.

f) *For each positive integer n,*
$\Lambda_{2n-1}(V)/\Lambda_{2n}(V) \cong \mathfrak{F}^+ \oplus \mathfrak{F}^+$.

It will be the main task of the remainder of this paper to prove the following result.

4.7. THEOREM. *Let V be a 3-dimensional anisotropic quadratic space over a nondyadic local field. Let Γ be a non-trivial normal subgroup of $\Omega(V)$. Let n be the largest integer such that $\Gamma \subseteq \Lambda_n(V)$. Then $\Lambda_{n+1}(V) \subset \Gamma$.*

Thus every non-trivial normal subgroup of $\Omega(V)$ is "sandwiched" between two successive $\Lambda(V)$'s.

However, instead of proving the above theorem, we will prove the following theorem which is clearly equivalent by virtue of 3.4 and 3.6.

4.8. THEOREM. *Let A be a quaternion division algebra over a nondyadic local field. Let $G = \{t \in A \mid N(t) = 1\}$. Suppose C is a normal subgroup of G which strictly contains $\{-1, 1\}$. Let n be the largest integer such that $C \subseteq G_n$. Then $G_{n+1} \subset C$.*

The proof of this theorem will be given in § 7.

5. Some results on quaternion algebras. In this section A denotes a quaternion division algebra over F and, as usual, $G = \{t \in A \mid N(t) = 1\}$. We now collect several results of the author that will be used repeatedly in the sequel.

5.1. *Let* $t \in G$, $t \neq -1$ *or* 1. *Let* $u \in G$, $u \neq t$ *or* t^{-1} *and suppose* $S(t) = S(u)$. *Then there exists* $x \in G$ *such that* $xtx^{-1} = u$ *if and only if*

$$(N(u - t^{-1}), S(t)^2 - 4)$$

splits.

Note that the hypothesis implies neither $N(u - t^{-1})$ nor $S(t)^2 - 4$ is zero. For a proof, see [11], Theorem 2. Although it is only stated there for pure quaternions, the translation to the above condition is trivial.

5.2. *Same hypotheses as in* 5.1. *Then exactly one of the following obtains.*

a) *There exists* $x \in G$ *such that*

$$xtx^{-1} = u.$$

b) *There exists* $y \in A$, $Ny = -1$, *such that*

$$yt^{-1}y^{-1} = u.$$

Proof. Immediate from [11], Theorem 4 since A is a division algebra. q. e. d.

5.3. *Let* $u \in G$, $u \neq -1$ *or* 1. *Suppose* $\alpha \in F$. *There exist* $t_1, t_2 \in G$ *such that*

a) $u = t_1 t_2$

b) $S(t_\nu) = \alpha$ *for* $\nu = 1, 2$

if and only if

$$(S(u)^2 - 4, S(u)^2 - 4 + \alpha^2(2 - S(u))) \cong A$$

provided $S(u)^2 - 4 + \alpha^2(2 - S(u)) \neq 0$.

Note that $u \neq -1$ or 1 implies $S(u)^2 - 4 \neq 0$ since A is a division algebra.

Proof. Immediate by [12], Theorem 2. q. e. d.

6. Generators for the groups G_n. As the title of this section suggests, we will determine a set of generators for G_n, $n \geqq 1$. Recall that (π, nonsquare unit) $\cong A$, our quaternion division algebra.

6.1. *Let* n *be a positive integer.* G_n *is generated by* -1 *and the set of all* $t \in G_n$ *for which*

$$(7) \qquad S(t) = 2(1+\pi^n\alpha)$$

for some unit α.

By 2.3 we note that necessarily 2α is a nonsquare if n is even.

Proof. Let $u \in G_n$ which we may suppose is neither -1 nor 1. Multiplying by -1 if necessary, we may suppose $\chi(u) = 1$. We shall prove that u is a product of elements whose traces have the desired form (7). It clearly suffices to consider the case in which $S(u) = 2(1+\pi^{n+l}\beta)$ where l is a positive integer and β is a unit. Note that necessarily 2β is a nonsquare if $n+l$ is even.

a) n even, l odd. Let α be a unit such that 2α is a nonsquare. Now $S(u)^2-4 \sim 2\pi\beta$ and $S(u)^2-4+4(1+\pi^n\alpha)^2(2-S(u)) \sim -\pi\alpha\beta$. Now $(2\pi\beta, -\pi\alpha\beta) \cong (\pi, 2\alpha) \cong A$. By 5.3 there exist $t_1, t_2 \in G$ such that $u = t_1t_2$ with

$$S(t_1) = S(t_2) = 2(1+\pi^n\alpha).$$

By 2.1, it follows that $t_1, t_2 \in G_n$ and we are done.

b) n even, l even. By 2.3 there exists $v \in G_n$ such that $S(v) = 2(1+\pi^{n+1})$. As in a), by employing 5.3 we prove the existence of $v_1, v_2 \in G_n$ such that $u = v_1v_2$ with $S(v_1) = S(v_2) = S(v)$. By case a) we may now write $v_1 = t_1t_2$, $v_2 = t_3t_4$ where the $t_\nu \in G_n$ and have traces of form (7). Thus $u = t_1t_2t_3t_4$ is expressed in the desired form.

c) n odd, l odd. Note that 2β is a nonsquare. Now $S(u)^2-4 \sim 2\beta$ and

$$S(u)^2-4+4(1+\pi^n)^2(2-S(u)) \sim -\pi\beta.$$

Thus $(2\beta, -\pi\beta) \cong (2\beta, \pi) \cong A$. By 5.3 there exist $t_1, t_2 \in G$ such that $u = t_1t_2$ with $S(t_1) = S(t_2) = 2(1+\pi^n)$. By 2.1, it follows that $t_1, t_2 \in G_n$ and we are done.

d) n odd, l even. Let α be a unit such that 2α is a nonsquare. By 2.3 there exists $v \in G_n$ such that $S(v) = 2(1+\pi^{n+1}\alpha)$. As in c) by employing 5.3 we prove the existence of $v_1, v_2 \in G_n$ such that $u = v_1v_2$ with $S(v_1) = S(v_2) = S(v)$. By case c) we may now write $v_1 = t_1t_2$, $v_2 = t_3t_4$ where the $t_\nu \in G$ and have traces of the form (7). Thus $u = t_1t_2t_3t_4$ is expressed in the desired form. q. e. d.

7. Proof of Theorem 4.8. We are now ready to give the proof of Theorem 4.8.

7.1. *Suppose that the largest integer n such that $C \subseteq G_n$ is even and positive, say $n = 2m > 0$. Then $G_{n+1} \subset C$.*

Proof. By the definition of n there exists $t \in C$ such that

$$S(t) = 2(1 + \pi^{2m}\alpha)$$

where α is a unit. Necessarily 2α is a nonsquare. By 6.1 it is enough to show that any $u \in G_{2m+1}$ with $S(u)$ of the form $2(1 + \pi^{2m+1}\beta)$, β a unit, is in C. Now part a) of the proof of 6.1 already gives

$$u = t_1 t_2 \text{ where } S(t_1) = S(t_2) = S(t)$$

and $t_1, t_2 \in G_{2m}$. If we can show that t_1 and $t_2 \in C$ we will be done. If t_ν ($\nu = 1$ or 2) $= t$ or t^{-1} there is nothing to prove so we may assume $t_\nu \neq t$ or t^{-1}. We are thus in a position to use 5.1. Now $S(t)^2 - 4 = 4\pi^{2m}(2\alpha + \pi^{2m}\alpha^2)$. The identity $N(t_\nu - t^{-1}) + N(t_\nu - t) = 4 - S(t)^2$ thus yields either

$$\operatorname{ord} N(t_\nu - t^{-1}) = 2m \text{ or } \operatorname{ord} N(t_\nu - t) = 2m.$$

a) $\operatorname{ord} N(t_\nu - t^{-1}) = 2m$. Thus

$$N(t_\nu - t^{-1}) = \pi^{2n}\gamma, \ \gamma \text{ a unit.}$$

Then $(N(t_\nu - t^{-1}), S(t)^2 - 4) = (\pi^{2m}\gamma, 4\pi^{2m}(2\alpha + \pi^{2m}\alpha^2)) \cong (\gamma, 2\alpha)$. Since (unit, unit) splits over a nondyadic local field (see [10], 63:12), 5.1 yields $x \in G$ such that $t_\nu = x t x^{-1}$. Thus $t_\nu \in C$.

b) $\operatorname{ord} N(t_\nu - t) = 2m$. Then 5.1 yields $x \in G$ such that $t_\nu = x t^{-1} x^{-1}$. Again $t_\nu \in C$ and we are done. q. e. d.

7.2. *Suppose that the largest integer n such that $C \subseteq G_n$ is odd, say $n = 2m - 1$. Then $G_{n+1} \subset C$.*

Proof. By the definition of n there exists $t \in C$ such that

$$S(t) = 2(1 + \pi^{2m-1}\alpha)$$

where α is a unit. By 6.1 it is enough to show that any $u \in G_{2m}$ with $S(u)$ of the form $2(1 + \pi^{2m}\beta)$, β a unit, is in C. Again note that 2β is a nonsquare. Now part c) of the proof of 6.1 already gives

$$u = t_1 t_2 \text{ where } S(t_1) = S(t_2) = S(t)$$

and $t_1, t_2 \in G_{2m-1}$. If we can show that t_1 and $t_2 \in C$ we will be done. If t_ν ($\nu = 1$ or 2) $= t$ or t^{-1} there is nothing to prove so we may assume $t_\nu \neq t$ or t^{-1}.

a) -1 is a square in F. By 5.2 either $t_\nu = xtx^{-1}$ with $x \in G$ in which case $t_\nu \in C$ or $t_\nu = yt^{-1}y^{-1}$ with $N(y) = -1$. Write $-1 = \zeta^2$, $\zeta \in F$ and $z = \zeta y$. Then $z \in G$ and $t_\nu = zt^{-1}z^{-1}$. Again $t_\nu \in C$.

b) -1 is a nonsquare of F. In this case we show directly that $u \in C$. Now $t_1 \neq t_2$ or t_2^{-1}. This is clear in the case of t_2^{-1} for $u = t_1 t_2 \neq -1$. If $t_1 = t_2$ then $u = t_1 t_2 = t_2{}^2 = S(t_2)t_2 - 1$ whence $S(u) = S(t_2)^2 - 2$ has the form $2(1 + \pi^{2m-1}\gamma)$, γ a unit, contrary to the choice of u. As $t_1 \neq t_2$ or t_2^{-1} we are able to use 5.1. Now

$$N(t_1 - t_2^{-1}) = 2 - S(t_1 t_2) = 2 - S(u) = -2\pi^{2m}\beta \sim -2\beta.$$

And $S(t_2)^2 - 4 \sim 2\pi\alpha$. Thus

$$(N(t_1 - t_2^{-1}), S(t_2)^2 - 4) \cong (-2\beta, 2\pi\alpha) \cong (-1, \pi)(2\beta, \pi).$$

Since both -1 and 2β are nonsquares, we see that $(N(t_1 - t_2^{-1}), S(t_2)^2 - 4)$ splits. By 5.1, there exists $x \in G$ such that $t_1 = xt_2x^{-1}$. Thus $u = t_1 t_2 = xt_2x^{-1}t_2$. By 5.2 either $t_2 = ztz^{-1}$ with $z \in G$ in which case $u \in C$ and we are done or else $t_2 = zt^{-1}z^{-1}$ with $Nz = -1$. In this case

$$\begin{aligned} u &= xzt^{-1}z^{-1}x^{-1}zt^{-1}z^{-1} \\ &= z\{[(z^{-1}xz)t^{-1}(z^{-1}xz)^{-1}]t\}z^{-1} = zvz^{-1} \text{ with } v \in C. \end{aligned}$$

If $u = v$ or v^{-1} then $u \in C$ and we are done.
So we may assume $u \neq v$ or v^{-1}. Since $S(u) = S(zvz^{-1}) = S(v)$ we are in a position to use 5.1.

$$S(v)^2 - 4 = S(u)^2 - 4 = 4\pi^{2m}(2\beta + \pi^{2m}\beta^2).$$

The identity $N(u - v^{-1}) + N(u - v) = 4 - S(v)^2$ thus yields either

$$\operatorname{ord} N(u - v^{-1}) = 2m \text{ or } \operatorname{ord} N(u - v) = 2m.$$

Proceeding as in the proof of 7.1 we conclude that either there exists $x \in G$ such that $u = xvx^{-1}$ or there exists $y \in G$ such that $u = yv^{-1}y^{-1}$. In either event, $u \in C$. q. e. d.

We now treat the only remaining case.

7.3. *Suppose C is not contained in G_1. Then $G_1 \subset C$.*

Proof. By 6.1 we need only show that every $u \in G_1$ with $S(u) = 2(1 + \pi\beta)$, β a unit, is in C. By hypothesis, there exists $t \in C$ such that $S(t) \not\equiv 2$ or $-2 \pmod{\mathfrak{p}}$. Or, equivalently, $S(t)^2 - 4$ is a unit. Writing $t = \frac{1}{2}S(t) + x$, x pure, we have $4 - S(t)^2 = N(2x)$. By 2.2, $S(t)^2 - 4$ is a nonsquare. Now $S(u)^2 - 4 \sim 2\pi\beta$ and $S(u)^2 - 4 + S(t)^2(2 - S(u)) \sim 2\pi\beta(4 - S(t)^2)$.

And $(2\pi\beta, 2\pi\beta(4-S(t)^2)) \cong (\pi, S(t)^2-4) \cong A$. By 5.3, there exist $t_1, t_2 \in G$ such that $u=t_1t_2$ with $S(t_1)=S(t_2)=S(t)$. If we can show that t_1 and $t_2 \in C$ we will be done. If t_1 and $t_2=t$ or t^{-1} there is nothing to prove so we may assume that $t_\nu \neq t$ or t^{-1} for $\nu=1,2$. We are in a position to apply 5.1. The identity $N(t_\nu-t^{-1})+N(t_\nu-t)=4-S(t)^2$ implies at least one of $N(t_\nu-t^{-1})$, $N(t_\nu-t)$ is a unit. Hence at least one of $(N(t_\nu-t^{-1}), S(t)^2-4)$, $(N(t_\nu-t), S(t)^2-4)$ splits. By 5.1 either $t_\nu = xtx^{-1}$ or $xt^{-1}x^{-1}$ with $x \in G$. Hence $t_\nu \in C$ and we are done. q. e. d.

8. Determination of all normal subgroups of $\mathbf{\Omega}(V)$. In this section we shall determine all the normal subgroups of $\Omega(V)$. Let Γ be a non-trivial normal subgroup of $\Omega(V)$. By Theorem 4.7 there exists a non-negative integer n such that $\Lambda_{n+1}(V) \subset \Gamma \subseteq \Lambda_n(V)$. Thus we need only determine which subgroups "sandwiched" between successive Λ's are normal in $\Omega(V)$.

8.1. THEOREM. *Let n be a non-negative integer. Let Γ be a subgroup of $\Omega(V)$ satisfying*

$$\Lambda_{2n+1}(V) \subset \Gamma \subseteq \Lambda_{2n}(V).$$

Then Γ is normal in $\Omega(V)$.

Proof. If $n=0$ the result immediately follows from the observations $\Omega(V)=\Lambda_0(V)$ and $\Lambda_0(V)/\Lambda_1(V)$ is abelian. So we may assume $n>0$. By virtue of 3.4 and 3.6 it suffices to prove that every subgroup C of G satisfying $G_{2n+1} \subset C \subseteq G_{2n}$ is normal in G. We utilize the epimorphism g defined in the proof of 4.1. We immediately have C is normal in G if and only if $g(xCx^{-1})=g(C)$ for all $x \in G$. We now verify this condition. Let $t \in C$. If $t \in G_{2n+1}$ the result is immediate so we may assume that $t \notin G_{2n+1}$. By 2.3 we may write

$$t=(\chi(t)+\pi^{2n}\alpha)1+\pi^n(t_1i+t_2j+t_3ij).$$

Set $x=x_01+x_1i+x_2j+x_3ij$. By direct computation we have $g(xtx^{-1}) = \rho(x_0^2-\epsilon x_1^2)g(t)$. But $Nx=1$ implies $\rho(x_0^2-\epsilon x_1^2)=1$ and we are done. q. e. d.

We need two lemmas.

8.2. *Let p be an odd prime, k and l positive integers. Suppose $k<2l$. Then the condition $(p^l+1)/2$ divides p^k-1 implies $p=3$ and $k=l=1$.*

Proof. By hypothesis there exists an integer r such that

$$(*) \qquad 2(p^k-1)=r(p^l+1).$$

1) Suppose $k \leqq l$. Then $p^k \leqq p^l$ whence $2(p^k - 1) < 2(p^l + 1)$. Then (*) implies $r = 1$ and we have $2p^k = p^l + 3$. Thus $p = 3$. And $2 \cdot 3^{k-1} = 3^{l-1} + 1$. We must have $k = 1$. Then $3^{l-1} = 1$ and we have $l = 1$.

2) Suppose $l < k$. We may write $2[p^{2l} - 1 - p^k(p^{2l-k} - 1)] = r(p^l + 1)$. Then $2(p^l - 1) - r = 2p^k \left(\dfrac{p^{2l-k} - 1}{p^l + 1}\right)$. Since $2(p^l - 1) - r$ is an integer and $p^l + 1$ is relatively prime to p^k we must have $(p^l + 1)/2$ divides $p^{2l-k} - 1$. But $2l - k < l$. By case 1) this is impossible. q. e. d.

8. 3. *Let* $\mathfrak{V}$ *be a regular 2-dimensional quadratic space over a finite field* $\mathfrak{F}$. *Let* $\mathfrak{G}$ *be a subgroup of* $O^+(\mathfrak{V})$ *and* $\mathfrak{W}$ *an additive subgroup of* $\mathfrak{V}$. *Suppose* $g(w) \in \mathfrak{W}$ *for each* $g \in \mathfrak{G}$ *and each* $w \in \mathfrak{W}$. *If* $|X|$ *denotes the cardinality of the set* X, *then* $|\mathfrak{G}|$ *divides* $|\mathfrak{W}| - 1$.

Proof. We may assume $|\mathfrak{W}| > 1$. Let $x \in \mathfrak{W}$, $x \neq 0$. Then $g_1(x) = g_2(x)$ for $g_1, g_2 \in \mathfrak{G}$ implies $g_1 = g_2$ (see [10], 43:2). Let $n = |\mathfrak{G}|$ and $\mathfrak{G} = \{g_1, \cdots, g_n\}$. Then the vectors $g_1(x), \cdots, g_n(x)$ are distinct and non-zero. If they exhaust the non-zero elements of $\mathfrak{W}$ we are done. If not there exists $y \neq 0$ in $\mathfrak{W}$ distinct from all the $g_i(x)$. Then $g_i(x) \neq g_j(y)$ for $i, j = 1, \cdots, n$. If the $2n$ vectors $g_i(x), g_j(y)$ $(i, j = 1, \cdots, n)$ exhaust the non-zero elements of $\mathfrak{W}$ we are done. If not we continue until all non-zero elements are exhausted. If this happens at the m-th stage, then $nm = |\mathfrak{W}| - 1$ and we are done. q. e. d.

8. 4. THEOREM. *Let* n *be a positive integer. Then there is no normal subgroup* Γ *of* $\Omega(V)$ *satisfying*

$$\Lambda_{2n}(V) \subset \Gamma \subset \Lambda_{2n-1}(V) \tag{*}$$

with exactly one exception: the residue class field of $F = \boldsymbol{F}_3$, *the field of* 3 *elements. In this case every subgroup* Γ *satisfying* (*) (*there are* 4 *such*) *is normal in* $\Omega(V)$.

Proof. By virtue of 3. 4 and 3. 6 it suffices to prove that no subgroup C of G satisfying $G_{2n} \subset C \subset G_{2n-1}$ is normal in G except when $\mathfrak{F} = \boldsymbol{F}_3$ when all are normal. We utilize the epimorphism h defined in the proof of 4. 2. We immediately have C is normal in G if and only if

$$h(xCx^{-1}) = h(C) \text{ for all } x \in G. \tag{8}$$

We now examine this condition.

Let $t \in C$. Then $t \in G_{2n-1}$ and we may write

$$t = (\chi(t) + \pi^{2n-1}t_0)1 + \pi^{n-1}(t_1 i + t_2 j + t_3 ij).$$

Set $x = x_0 1 + x_1 i + x_2 j + x_3 ij$. Let $\rho(x_0) = \xi_0$, $\rho(x_1) = \xi_1$ and $\rho(\epsilon) = e$. (Note that $Nx = 1$ implies $\xi_0^2 - e\xi_1^2 = 1$). Then, by direct computation, we have

$$h(xtx^{-1}) = h(t)\begin{pmatrix} \xi_0^2 + e\xi_1^2 & 2\xi_0\xi_1 \\ 2e\xi_0\xi_1 & \xi_0^2 + e\xi_1^2 \end{pmatrix}. \tag{9}$$

We now interpret this geometrically. $\mathfrak{F}^+ \oplus \mathfrak{F}^+$ is a 2-dimensional vector space $\mathfrak{V}$ over $\mathfrak{F}$ with base $E_1 = (1,0)$, $E_2 = (0,1)$. We define a bilinear form $\mathfrak{B}$ on $\mathfrak{V}$ by setting $\mathfrak{B}(E_1, E_1) = 1$, $\mathfrak{B}(E_2, E_2) = -e$ and $\mathfrak{B}(E_1, E_2) = 0$. Thus $\mathfrak{V}$ has the structure of a quadratic space and, by the Local Square Theorem, it is anisotropic. Now using the spinorial representation for $O^+(\mathfrak{V})$ we see that the matrix in (9) is the generic element of the commutator subgroup $\Omega(\mathfrak{V})$ of $O(\mathfrak{V})$ since $\xi_0^2 - e\xi_1^2 = 1$. Hence, by (9), condition (8) can be restated as follows:

$$\Sigma(X) \in h(C) \text{ for all } X \in h(C) \text{ and all } \Sigma \in \Omega(\mathfrak{V}). \tag{10}$$

By 8.3 we must have

$$|\Omega(\mathfrak{V})| \text{ divides } |h(C)| - 1. \tag{11}$$

Suppose $|F| = p^l$. Then $|h(C)| = p^k$ where $k < 2l$ and $|\Omega(\mathfrak{V})| = (p^l + 1)/2$ (see [1], pp. 147 and 210). By 8.2, (11) can only happen if $p = 3$ and $l = 1$, i.e., when $\mathfrak{F} = \boldsymbol{F}_3$. Thus we have proved that no normal subgroups Γ exist unless $\mathfrak{F} = \boldsymbol{F}_3$. However when $\mathfrak{F} = \boldsymbol{F}_3$, $\Omega(\mathfrak{V})$ consists of the identity and its negative and (10) holds trivially. Thus every subgroup Γ in this case is normal. These subgroups are, of course, in bijective correspondence with the additive subgroups of $\mathfrak{V}$ distinct from 0 and $\mathfrak{V}$, that is, with the 1-dimensional subspaces of $\mathfrak{V}$. There are 4 such. q. e. d.

UNIVERSITY OF NOTRE DAME,
NOTRE DAME, INDIANA.

REFERENCES.

[1] E. Artin, *Geometric Algebra*, New York, 1957.
[2] ———, *Algebraic numbers and algebraic functions I*, Princeton University, New York University, 1950-1951.

[3] J. Dieudonné, *Sur les groupes classiques,* Paris, 1948.

[4] ———, "Sur les groupes orthogonaux rationnels à trois et quatre variables," *Comptes Rendus Hebdomadaires des Séances de l'Academie des Sciences* (Paris), vol. 233 (1951), pp. 541-543.

[5] M. Eichler. *Quadratische Formen und orthogonale Gruppen,* Berlin, 1952.

[6] A. Hattori, "On the multiplicative group of simple algebras and orthogonal groups of three dimensions," *Journal of the Mathematical Society of Japan,* vol. 4 (1952), pp. 205-217.

[7] N. Jacobson, *Structure of Rings,* American Mathematical Society Colloquium Publications, vol. 37 (1956).

[8] B. W. Jones, *The Arithmetic Theory of Quadratic Forms,* Carus Monographs, No. 10, Mathematical Association of America, 1950.

[9] R. Lipschitz, "Correspondence," *Annals of Mathematics,* vol. 69 (1959), pp. 247-251.

[10] O. T. O'Meara, *Introduction to Quadratic Forms,* Berlin, 1963.

[11] B. Pollak, "The equation $\bar{t}at = b$ in a quaternion algebra," *Duke Mathematical Journal,* vol. 27 (1960), pp. 261-272.

[12] ———, "Conjugacy in the real three dimensional orthogonal groups," *Illinois Journal of Mathematics,* vol. 8 (1964), pp. 5-13.

SINGULAR INTEGRALS AND BOUNDARY VALUE PROBLEMS.

By R. T. Seeley.*

Introduction. The present article begins to develop the approach to boundary problems introduced by A. P. Calderón [6], and exploited for certain first order systems in [11]. The inherent connection between boundary problems and singular integrals on the boundary is well known in the case of the Cauchy-Riemann equations, which we discuss briefly in Section 1. The main theorem of this paper generalizes this situation to an arbitrary system. Its corollaries contain many of the published results on elliptic boundary problems, and we feel that this approach illuminates these results. The corollaries will be presented in subsequent papers.

The results are given in the framework of the spaces L_s^p of Bessel potentials, and their restrictions $B_{s-1/p}^p$. Thus we rely on results of Calderón [4], Stein [14], and Taibleson [15], and use the L^p theory of Calderón-Zygmund operators (alias pseudo-differential operators) of [12]. However, the case $p = 2$ is the most important, and the paper may be read with the background of [8] or [11] if in all statements L_s^p and $B_{s-1/p}^p$ are replaced by H^s and $H^{s-1/2}$.

The main result can be described as follows. Let E be a vector bundle over a manifold G_+ with boundary X, and let A be an elliptic operator of order m from sections of E to sections of some other bundle F. Define $L_s^p(E)$ to be those sections of E having an extension in $L_s^p(2E)$, where $2E$ is an extension of E over the double of G_+. Let R denote the map of $u \in C^\infty(E)$ onto its Cauchy data of order $< m$ on X. Let $N_0(A) = \{u \in C^\infty(E) : Au = 0$ and $Ru = 0\}$, and $K(A,s,p) = \{u \in L_s^p(E) : Au = 0$ in the interior of $G_+\}$.

THEOREM. *For all real s and $1 < p < \infty$, R is continuous from $K(A,s,p)$ to $\oplus_1^m B_{s-1/p-j+1}^p(E \mid X) = B_{s,m}^p$. There is a Calderón-Zygmund operator P^+ which projects $B_{s,m}^p$ onto $R(K(A,s,p))$ for each s. The top-order symbol of P^+ is the projection onto certain generalized eigenspaces of an operator derived from the symbol of A. There is a continuous map P of $B_{s,m}^p$ into*

Received July 21, 1965.

* This work supported in part by a grant from the National Science Foundation, at Brandeis University.

$K(A,s,p)$, whose range has $N_0(A)$ as a linear complement, and such that $RP = P^+$.

An application of this result (not discussed in this article) is the following. Let $A_{\max}$ be the unbounded operator on $L^p(E)$ whose domain is $D_{\max} = \{u \in L^p(E) : Au \in L^p(E)\}$; and set $A_{\max}u = Au$. Then the restriction map R is a continuous map onto those m-tuples f of distributions such that $P^+f \in B_{0,m}{}^p$ and $f - P^+f \in B_{m,m}{}^p$. Call a realization of A in L^p any extension of $A_{\min}$, where the domain of $A_{\min}$ is $D_{\min} =$ the closure in $L_m{}^p(E)$ of the C^∞ sections vanishing near X. Then the realizations of A in L^p are in one-one correspondence with the subspaces of $P^+(B_{0,m}{}^p) \oplus (I - P^+)(B_{m,m}{}^p)$; the correspondence is achieved by applying the restriction map R to the domain of the realization. The realization is closed if and only if the subspace is closed, and the gap between the graphs of the closed realizations is equivalent to the gap between the subspaces.

The plan of the paper is as follows. § 1 describes the special case of the Cauchy-Riemann equations in the unit disk, and indicates the approach to be taken for a general elliptic system. § 2 reviews the necessary results for the spaces $L_s{}^p$ and $B_s{}^p$, and § 3 discusses CZO's (alias pseudo-differential operators) on compact manifolds without boundary. § 4 states a special case of the main result of the paper, in Theorems 5 and 6, concerning a basic projection P^+ and associated "multiple layer" potentials P. § 5 constructs P^+, proves it is a CZO and establishes the main lemmas for the remaining proofs. § 6 defines the multiple layer potential P, proves Theorem 5, and describes the symbol of P^+. § 7 proves Theorem 6, and § 8 extends the results to the general case, stated as Theorem 7.

The underlying idea of the proof is indicated in [6]: if A is an elliptic operator of order m on a manifold G_+ with boundary X, then the Cauchy data on X of order $< m$ of functions in G_+ splits naturally into two complementary spaces. One of these is the space of Cauchy data of solutions u in G_+ of $Au = 0$; and if A extends to an elliptic operator A on the double G of G_+, then the complementary space is essentially the space of Cauchy data of solutions v of $Av = 0$ in $G - G_+$.

The author is indebted to A. P. Calderón for several discussions relating to these results.

1. The Cauchy-Riemann equations in the unit disk. Consider the operator $A = \partial/\partial\bar{z} = 1/2(\partial/\partial x + i\partial/\partial y)$. Then $Au = 0$ means u is holomorphic. Let D denote the disk $\{x^2 + y^2 < 1\}$, and $C =$ boundary of D. Let H_+ be the space of functions u, holomorphic in D, and such that u and its first derivatives are in $L^2(D)$. Let H_- be the space of functions v holomorphic in the complement of D, vanishing at ∞, and such that v and its first derivatives are in L^2 on any compact subset of the complement of D. Finally, let $H^{1/2}(C)$ be the class of functions $f=\sum_{-\infty}^{\infty} a_n e^{in\theta}$ with $\sum(1+|n|)|a_n|^2<\infty$. Then for $|z|<1$ let $Pf(z)=\sum_{0}^{\infty} a_n z^n$, and for $|z|>1$ set $Pf(z)=\sum_{-\infty}^{-1} a_n z^n$, when $f=\sum_{-\infty}^{\infty} a_n e^{in\theta}$. Then $f\rightarrow(Pf$ restricted to $D)$ maps $H^{1/2}(C)$ onto H_+, and $f\rightarrow(Pf$ restricted to $\{x^2+y^2>1\})$ map $H^{1/2}(C)$ onto H_-. The boundary values $P^+f(e^{i\theta})=\sum_{0}^{\infty} a_n e^{in\theta}$ and $P^-f(e^{i\theta})=\sum_{-\infty}^{-1} a_n e^{in\theta}$ yield projections $P^{\pm}$ dividing $H^{1/2}(C)$ into complementary spaces, the spaces of boundary values of functions in H_+ or H_- respectively. P^+ and P^- are the basic singular integral operators on the unit circle, and their connection with holomorphic functions is widely exploited; see e. g., Zygmund [16], Chapter 7.

To make an application in a standard boundary problem, consider the system $A'=\tfrac{1}{2}\begin{pmatrix}\partial/\partial x & -\partial/\partial y\\ \partial/\partial y & \partial/\partial x\end{pmatrix}$, acting on pairs of real functions. If we let a complex u correspond to the pair $(\mathrm{Re}(u),\mathrm{Im}(u))$, then A corresponds to A', (complex) $H^{1/2}(C)$ to (real) $H^{1/2}(C)\oplus H^{1/2}(C)$, and $H_{\pm}$ to spaces $H_{\pm}'$, P to P', $P^{\pm}$ to $P^{\pm\prime}$. If now B is a bounded linear map of $H^{1/2}(C)\oplus H^{1/2}(C)$ into $H^{1/2}(C)$, then we say that B is well-posed for A' on D if

(i) $(BP^{+\prime})^{-1}(0)/(P^{+\prime})^{-1}(0)$ and

(ii) $H^{1/2}(C)/BP^{+\prime}(H^{1/2}(C)\oplus H^{1/2}(C))$

are both finite dimensional. These conditions are satisfied, e. g., if $B(f,g)=f$. When B is well-posed, it is easy to check the following facts (a) and (b).

(a) Consider the map $(A',B):(u,v)\rightarrow[A'(u,v),B(u\mid C,v\mid C)]$, defined when u and v and their first derivatives are in $L^2(D)$. Then (A',B) has finite dimensional kernel and co-kernel, as map into

$$L^2(D)\oplus L^2(D)\oplus H^{1/2}(C).$$

(b) Consider the map $A'_B:(u,v)\rightarrow A'(u,v)$ with domain all pairs (u,v) in $L^2(D)$ together with their first derivatives, and such that $B(u\mid C,v\mid C)=0$. Then A'_B has finite dimensional kernel and cokernel as a map into $L^2(D)\oplus L^2(D)$.

To illustrate the methods, we show that the cokernel of (A',B) has finite dimension. Suppose a triple $(u',v';f)$ is given in $L^2(D)\oplus L^2(D)\oplus H^{1/2}(C)$.

Then by extending u' and v' over the whole plane so that they have compact support and mean value zero, we obtain via Fourier transforms a pair (u_1, v_1), in $L^2(D)$ together with their first derivatives, with $A'(u_1, v_1) = (u', v')$. From general results on restrictions (e. g. Theorem 1 below), $u_1 \mid C$ and $v_1 \mid C$ are in $H^{1/2}(C)$; and from the finite dimensionality of (ii), $f - B(u_1 \mid C, v_1 \mid C)$ is in the range of $BP^{+\prime}$ if it is orthogonal to a finite set of functions. If these conditions are satisfied, then $f - B(u_1 \mid C, v_1 \mid C) = BP^{+\prime}(g, h)$ for some pair (g, h), and setting $(u_2, v_2) = P'(g, h)$ we have $A'(u_2, v_2) = 0$, and $(u_2 \mid C, v_2 \mid C) = P^{+\prime}(g, h)$, so that $(A', B)(u_1 + u_2, v_1 + v_2) = (u', v'; f)$.

Conversely either (a) or (b) implies the finite dimensionality of (i) and (ii).

The general approach of Calderón is to find and analyze the corresponding projections $P^{\pm}$ for an arbitrary elliptic system. He announced [6] that the projections are singular integral opeerators, from which the conditions corresponding to (i) and (ii) above can be checked algebraically when B is likewise a singular integral operator. In the rest of this paper we construct the generalizations of the P and P^+ introduced above, show P^+ to be a CZO as in [8], [11] or [12], and discuss the behavior of P on various function spaces modeled on L^p, $1 < p < \infty$.

2. Distribution spaces. The results for boundary problems will be stated in terms of two scales of spaces, the "Bessel potential" spaces $L_s{}^p$ of [4], and the Besov spaces $B_s{}^p$ which are included in the results of [15], where they are denoted $\Lambda(s, p, p)$. Further references for results of this nature are given in those articles. Here we are concerned mainly with restrictions, extensions and interpolation.

For notations, $x = (x_1, \cdots, x_n)$ denotes a point in R^n, $\langle x, y\rangle = \sum x_j y_j$, $|x|^2 = \langle x, x\rangle$. When $\alpha = (\alpha_1, \cdots, \alpha_n)$ is an n-tuple of integers $\geqq 0$, then $|\alpha| = \sum \alpha_j$, $\alpha! = \prod_1^n (\alpha_j)!$, $x^\alpha = x_1{}^{\alpha_1} \cdots x_n{}^{\alpha_n}$. Further, $D_j = -i\partial/\partial x_j$, $D = (D_1, \cdots, D_n)$, $D^\alpha = D_1{}^{\alpha_1} \cdots D_n{}^{\alpha_n}$. If f is integrable,

$$f^\wedge(\xi) = \int e^{-i\langle x, \xi\rangle} f(x)\, dx.$$

$\mathcal{S}$ is the space of f in $C^\infty(R^n)$ such that $x^\alpha D^\beta f(x)$ is bounded for all α and β; the topology of $\mathcal{S}$ is defined by the sequence of norms

$$\| f \|_k = \sup\{(1 + |x|^k |D^\beta f(x)| \text{ for } |\beta| \leqq k,\ x \text{ in } R^n\}.$$

$\mathcal{S}'$ is the dual of $\mathcal{S}$. The transpose of the Fourier transform on $\mathcal{S}$ is an automorphism of $\mathcal{S}'$, and is likewise denoted by $^\wedge$.

Definition 1. For s real, J^s is the automorphism of $\mathcal{S}'$ defined by $(J^s f)^\wedge = G^s f^\wedge$, where $G^s(\xi) = (1+|\xi|^2)^{-s/2}$.

$L^p(R^n)$ is the space of measurable functions ϕ such that $|\phi|^p$ is integrable. $L_0^p(R^n)$ is the space of distributions f (members of $\mathcal{S}'$) of the form $f(\psi) = \int \psi\phi$ for some ϕ in L^p and all ψ in $\mathcal{S}$. Then $\|f\|_{L_0^p} = (\int|\phi|^p)^{1/p}$. The function ϕ is called the density of f.

In the rest of this paper, *p and q will denote reals strictly between* 1 *and* ∞.

Definition 2. $L_s^p(R^n)$ is the isometric image of $L_0^p(R^n)$ *under* J^s.

In [4] it is shown that f is in $L_k^p(R^n)$ (k integer $\geqq 0$) if and only if there are functions ϕ_α ($|\alpha| \leqq k$) in $L^p(R^n)$ such that $(D^\alpha f)(\psi) = \int \psi\phi_\alpha$ for all ψ in $\mathcal{S}$. Then an equivalent norm on L_k^p is $\sum \|\phi_\alpha\|_{L^p}$. Also L_{-k}^p consists of the distributions of the form $f = \sum_{|\alpha|\leqq k} D^\alpha \phi_\alpha$, ϕ_α in L_0^p.

Definition 3. $B_1(R^n)$ is the space of distributions f in L_0^p whose density ϕ satisfies.

$$(1) \qquad \int\int |\phi(x+y) + \phi(x-y) - 2\phi(x)|^p \, |y|^{-n-p} \, dy dx < \infty.$$

For s real, $B_s^p(R^n)$ is defined as the image of B_1^p under J^{s-1}.

It follows from [15], Theorem 1, that B_s^p may be defined for $0 < s < 2$ by replacing the condition (1) with

$$\int\int |\phi(x+y) + \phi(x-y) - 2\phi(x)|^p \, |y|^{-n-sp} \, dy dx < \infty.$$

It is known [15] that $B_s^p \subset L_s^p$ if and only if $p \leqq 2$; and $L_s^p \subset B_s^p$ if and only if $p \geqq 2$. In particular, B_0^p differs essentially from L^p if $p \neq 2$. In another sense, however, the difference is not great: when $s < t$, then $B_s^p \supset L_t^p$ and $L_s^p \supset B_t^p$. For other inclusions for varying values of p, see [4] and [15].

The following results for L_s^p are given in [4]. $L_{-s}^{p'}$ is the dual of L_s^p when $1/p + 1/p' = 1$. C^∞ functions with compact support are dense in L_s^p. D^α is continuous from L_s^p to $L_{s-|\alpha|}^p$, and in fact an equivalent norm in L_s^p is $\sum_{|\alpha|\leqq k} \|D^\alpha f\|_{L_{s-k}^p}$.

The same results with B in place of L are given in [15].

We work with the spaces L_s^p because they have the simple and familiar form given above when s is an integer $\geqq 0$, and because appropriate interpolation theorems holds for them. Theorem 1 below, due to Stein, shows that the spaces B_s^p then enter naturally in connection with boundary problems.

Let $R_{m,t}$ be the map from $\mathcal{S}(R^n)$ to $\mathcal{S}(R^{n-1}) \oplus \cdots \oplus \mathcal{S}(R^{n-1})$ given by $(R_{m,t}\phi)(x_2, \cdots, x_n)$

$$= \phi(t, x_2, \cdots, x_n) \oplus D_1\phi(t, x_2, \cdots, x_n) \oplus \cdots \oplus D_1^{m-1}\phi(t, x_2, \cdots, x_n).$$

The statement below is a trivial extension of Stein's result [14].

THEOREM 1. a) *For* $s - m + 1 - 1/p > 0$, $R_{m,t}$ *is bounded from* $L_s^p(R^n)$ *to* $\oplus_1^m B_{s-1/p-j+1}^p(R^{n-1})$, *with a bound independent of* t.

b) *There is an extension operator* $\mathcal{E}_m$, *continuous from*

$$\bigoplus_1^m B_{s-1/p-j+1}^p(R^{n-1})$$

to $L_s^p(R^n)$ *for all real* s *and* $1 < p < \infty$, *such that* $\mathcal{E}_m(\phi_1, \cdots, \phi_m)$ *is* C^∞ *for* $x_1 \neq 0$, *and* $R_{m,t}\mathcal{E}_m(\phi_1, \cdots, \phi_m) \to (\phi_1, \cdots, \phi_m)$ *in* $\bigoplus_1^m B_{s-1/p-j+1}^p$ *as* $t \to 0$.

Proof. Let $x' = (x_2, \cdots, x_n)$ denote a generic point in R^{n-1}. A proof of (a) is given in [14] for $m = 1$, and follows for general m because D_1^j is bounded from L_s^p to L_{s-j}^p.

The proof of (b) is based on the following result from [14]: *If* λ *is a* C^∞ *function on* R^1 *with compact support, and* ψ *is* C^∞ *on* R^{n-1} *with compact support, then for* g *in* $B_{s-1/p}^p$, $s > 1/p$, *the function*

$$\mathcal{E}\psi(g)(x_1, \cdots, x_n) = \lambda(x_1) \int_{R^{n-1}} g(x' - x_1 y')\psi(y')\,dy' \tag{2}$$

is in $L_s^p(R^n)$, *the map* $g \to \mathcal{E}\psi(g)$ *being continuous; and*

$$R_{1,t}\mathcal{E}\psi(g) \to \lambda(0)\left(\textstyle\int\psi\right)g \text{ as } t \to 0.$$

In order to obtain prescribed Cauchy data of order $< m$ we need a ψ, C^∞ with compact support, such that

$$\int A^{k-1}\psi(y')\,dy' = \delta_{ik}, \qquad k = 1, \cdots, m, \tag{3}$$

with $A\psi = \sum_{j=2}^{n} D_j(y_j\psi)$. Such a ψ exists, since the maps $\psi \to \int A^{k-1}\psi$ are linearly independent, the A^{k-1} being of different orders. Then (b) is established for $s > m + 1/p - 1$ by setting $\mathcal{E}\psi(g_1, \cdots, g_m) = \sum_1^m \phi_j/j!$, where $\phi_j(x_1, \cdots, x_n) = x_1^{j-1}\mathcal{E}\psi(g_j)(x_1, \cdots, x_n)$ with $\mathcal{E}\psi(g_j)$ as in (2) and λ chosen to equal 1 in a neighborhood of zero. Then from the result quoted above, ϕ_j is in $L_{s-j+1}^p(R^n)$ when g_j is in $B_{s+1-j-1/p}^p(R^{n-1})$. Also

$$\begin{aligned} &(D_{x'})^\alpha D_1^l\phi_j \\ &= \lambda(x_1)\sum_{k=0}^{l} C_k^l (D_1^k x_1^{j-1}) x_1^{-|\alpha|} \int g_j(x' - x_1 y')(D_{y'})^\alpha A^{l-k}\psi(y')\,dy' \\ &\qquad + \text{similar terms involving derivatives of } \lambda. \end{aligned} \tag{4}$$

This again lies in $L_{s-j+1}{}^p$ for $l + |\alpha| \leqq j - 1$, so in fact ϕ_j is in $L_s{}^p(R^n)$. From (3) and (4), $R_{m,t}(\sum \phi_j/j!) \to (g_1, \cdots, g_m)$ in $\overset{m}{\underset{1}{\oplus}} B_{s-1/p-j+1}{}^p(R^{n-1})$, which completes the proof in case $s > m + 1/p$.

For the remaining cases choose an integer k so that $2k + s > m + 1/p$, and apply the above result to

$$(1 - \sum_2^n \partial^2/\partial x_j{}^2)^k \mathcal{E}_\psi(J^{2k} g_1, \cdots, J^{2k} g_m) = \mathcal{E}_\psi(g_1, \cdots, g_m),$$

where $(J^{2k} g_j)^\wedge = (1 + \sum_2^n \xi_j{}^2)^{-k} g_j{}^\wedge$. This completes Theorem 1.

An immediate corollary of part (a) of the theorem is that the map $t \to R_{m,t}u$ is continuous from R^1 into $\overset{m}{\underset{1}{\oplus}} B_{s-1/p-j+1}{}^p(R^{n-1})$ when u is in $L_s{}^p(R^n)$ with $s - m + 1 - 1/p > 0$; for the map is continuous when u is C^∞ with compact support, and such u are dense in $L_s{}^p$. We also obtain the

COROLLARY 1.1. *Let k be an integer > 0, and u and v be in $L_k{}^p(R^n)$. If $R_{k,0}u = R_{k,0}v$, then the function w: $w =$ density of u when $x_1 > 0$, $w =$ density of v when $x_1 < 0$, is the density of an element of $L_k{}^p(R^n)$.*

In fact, when u' and v' are C^∞ with compact support, and ϕ is the same, then for $|\alpha| \leqq k$ $\int w D^\alpha \phi = (-1)^{\alpha_1}[\int_{x_1>0} \phi D^\alpha u' + \int_{x_1<0} \phi D^\alpha v'] +$ boundary terms involving $R_{k,0}u'$ and $R_{k,0}v'$. Letting $u' \to u$ and $v' \to v$ we have from Theorem 1 that the boundary terms converge; in fact their sum converges to zero because $R_{k,0}u = R_{k,0}v$. Thus $|\int w D^\alpha \phi| \leqq \text{const} \|\phi\|_{L^{p'}}$, $(1/p + 1/p' = 1)$, and w is in $L_k{}^p(R^n)$.

Now let $R_+{}^n$ be the half space $x_1 > 0$, $C_o{}^\infty(R_+{}^n)$ be the C^∞ functions with compact support in $R_+{}^n$, and for $s \geqq 0$, let $L_s{}^p(R_+{}^n)$ be those linear functionals on $C_o{}^\infty(R_+{}^n)$ of the form $\phi \to f(\phi)$ for an f in $L_s{}^p(R^n)$, i.e. essentially those functions f in $L^p(R_+{}^n)$ having an extension in $L_s{}^p(R^n)$.

COROLLARY 1.2. *For k an integer > 0 there is a continuous linear map* Ext: $L_k{}^p(R_+{}^n) \to L_k{}^p(R^n)$ *such that* $\text{Ext}(f) = f$ *in* $R_+{}^n$.

In fact, if $f \in L_k{}^p(R^n)$ and we set

$$v(x_1, \cdots, x_n) = \sum_0^{k-1} a_j f(-j x_1, x_2, \cdots, x_n)$$

where $\sum_0^{k-1} (-j)^m a_j = 1$ for $0 \leqq m < k$, then the function $w = f$ in $R_+{}^n$, $w = v$

in R_-^n, depends only on the values of f in R_+^n, and by Corollary 1.1 it depends continuously on f.

This section concludes with the statement of some interpolation theorems, and an application. We refer to [4] and [15] for the major part of the proof.

THEOREM 2. *Let A be a linear operator from $\mathcal{S}(R^n)$ to $\mathcal{S}'(R^m)$, s_i and t_j real, p_j and q_j strictly between* 1 *and* ∞, *for* $j=1,2$.

a) *There is a constant C depending on t_j, s_j, p_j, q_j, such that if* $\| Af\colon L_{t_j}^{q_j} \| \leqq B \| f\colon L_{s_j}^{p_j} \|$, *then* $\| Af\colon L_{t(\theta)}^{q(\theta)} \| \leqq CB \| f\colon L_{s(\theta)}^{p(\theta)} \|$ *for*

$$s(\theta)=(1-\theta)s_1+\theta s_2, \quad 1/p(\theta)=(1-\theta)/p_1+\theta/p_2, \qquad 0\leqq\theta\leqq 1,$$

and $t(\theta)$, $q(\theta)$ *given similarly.*

b) *The result* (a) *holds if L_s^p and L_t^q are replaced by B_s^p and B_t^q.*

c) *The result* (a) *holds if L_s^p is replaced by B_s^p.*

Proof. (a) is given in [4], and (b) in [15]. Then (c) follows from (a) and Theorem 1. In fact, let R denote the restriction: $L_{s+1/p}^p(R^{n+1}) \to B_s^p(R^n)$ $(s>0)$, and $\mathcal{E}$ denote its continuous right inverse. Choose $k<s_1$, $k<s_2$. Consider AJ^kR which, by assumption, is bounded from $L_{s_j+1/p_j-k}^{p_j}(R^{n+1}) \to L_{t_j}^{q_j}(R^m)$. Then (c) follows by applying (a) to AJ^kR and multiplying on the right by $\mathcal{E}J^{-k}$.

From Theorem 2 follows a result (not a sharp one) on the localizability of the spaces B_s^p and L_s^p.

COROLLARY 2.1. *Let ϕ have derivatives of order $\leqq k$ bounded and measurable on R^n. Then the map $f\to\phi f$ is bounded on L_s^p for $1<p<\infty$ and $|s|\leqq k$, and on B_s^p for $1<p<\infty$ and $1-1/p-k\leqq s\leqq k-1/p$.*

Proof. For L_k^p with k an integer $\geqq 0$, the result follows from the characterization of L_k^p as those f in L^p whose derivatives of order $\leqq k$ are in L^p. It holds for L_{-k}^p by duality, and for L_s^p $(-k\leqq s\leqq k)$ by interpolation.

For B_s^p with $s=k-1/p$ we extend ϕ to a $\phi^\sim$ on R^{n+1} with bounded derivatives of order $\leqq k$, and write $\phi f = R\phi^\sim\mathcal{E}f$ with R and $\mathcal{E}$ as in the proof of Theorem 2(c). The other values of s are obtained by duality and interpolation.

A similar argument shows these spaces are invariant under diffeomorphisms of R^n.

3. Calderón-Zygmund operators. The L^2 theory of these operators (here called CZO's) is given in [8], [9] and [11], and the L^p theory in [12].

A linear operator $A: \mathcal{S}(R^n) \to \mathcal{S}'(R^n)$ is of *order* $\leqq r$ if and only if there are constants C_{sp} such that $\| Af: L_s^p(R^n) \| \leqq C_{sp} \| f: L_{s+r}^p(R^n) \|$ for each real s and $1 < p < \infty$; and the same holds with B in place of L.

Suppose $a(x, \zeta)$ is a function on $R^n \times R^n$, $a(x, t\zeta) = t^r a(x, \zeta)$ for $t > 0$, and for each α, β and γ, $x^\alpha D_x^\beta D_\zeta^\gamma a(x, \zeta)$ is bounded for x in R^n and $|\zeta| = 1$. Then a is called a homogeneous symbol of degree r. If θ is a C^∞ function with $\theta = 0$ in a neighborhood of zero and $\theta = 1$ in a neighborhood of ∞, then the operator $\operatorname{Op}(a, \theta)$ given by

$$\operatorname{Op}(a, \theta) f(x) = (2\pi)^{-n} \int e^{i\langle x, \zeta \rangle} a(x, \zeta) \theta(\zeta) f^\wedge(\zeta)\, d\zeta$$

is called a homogeneous CZO of order r. Such an operator is in fact of order $\leqq r$. The necessary bounds for L_s^p are given in [12], Theorem 1; we recall the proof briefly, and indicate how it extends to B_s^p. The method, as in [7], is to expand $a(x, \zeta)$ in spherical harmonics, $a(x, \zeta) = \sum a_{kl}(x) S_{kl}(\zeta/|\zeta|) |\zeta|^r$. Corollary 2.1 of Theorem 2 above provides bounds for the multiplication by a_{kl}, and Mihlin's theorem provides bounds for the multiplication of Fourier transforms $f^\wedge(\zeta)$ by $S_{kl}(\zeta/|\zeta|)$. In fact, setting $(Tf)^\wedge(\zeta) = S_{kl}(\zeta/|\zeta|) f^\wedge(\zeta)$, we have that T is a translation invariant operator on L^p, and $\| Tf: L^p \| \leqq C \| f: L^p \|$. Then

$$\begin{aligned} \int\int | Tf(x+y) - 2Tf(x) + Tf(x-y) |^p \, |y|^{-n-p} \, dx dy & \\ \leqq C \int\int | f(x+y) - 2f(x) + f(x-y) | \, |y|^{-n-p} \, dx dy&, \end{aligned}$$

and T is bounded on B_1^p. Since also T commutes with the isomorphism $J^k: B_s \to B_{s+k}^p$, the bound for T on B_1^p also holds for B_s^p for all real s.

It is further shown in [12] that $\operatorname{Op}(a, \theta)$ is not of order $< r$, unless $a = 0$.

The operators we are about to define are less general than those considered in the references above, but they are the natural class to consider in the present situation. They could be called CZO's "of integral type."

Definition 4. A linear map $A: \mathcal{S}(R^n) \to \mathcal{S}'(R^n)$ is a CZO if and only if there is a sequence $\{a_j\}_{-\infty}^{\infty}$ of symbols with a_j homogeneous of degree j, $a_j = 0$ for j sufficiently positive, and such that $A - \sum_{j \geqq k} \operatorname{Op}(a_j, \theta)$ is of order $\leqq k - 1$, for each k. The *symbol* of A is the (formal) sum $\sigma(A)(x, \zeta) = \sum a_j(x, \zeta)$. If, for some integer k, $a_j = 0$ for all $j > k$, then $\sigma_k(A)$ is defined to be a_k. Otherwise $\sigma_k(A)$ is not defined.

In particular, a differential operator $A = \sum_{|\alpha| \leqq m} a_\alpha D^\alpha$ (a_α in $\mathcal{S}$) is a CZO, and $\sigma(A)$ is its characteristic polynomial, $\sigma_m(A)$ the top order terms of the characteristic polynomial.

Since $\operatorname{Op}(a_j, \theta)$ is of order $\leqq j$, the condition "$A - \sum_{j \geqq k} \operatorname{Op}(a_j, \theta)$ is of order $\leqq k-1$ for each k" may be replaced by the formally weaker condition: given real numbers s, t and $1 < p < \infty$, there is a k such that $A - \sum_{j \geqq k} Op(a_j, \theta)$ is continuous from L_s^p to L_t^p and from B_s^p to B_t^p. Using the inclusions stated after Definition 3 and the characterization of L_t^p for $t =$ integer, it suffices to prove that given integers $s \leqq 0$ and $t \geqq 0$ there is a k such that

$$|\int (D^\alpha \phi)(A D^\beta \psi - \sum_{j \geqq k} \operatorname{Op}(a_j, \theta) D^\beta \psi | \leqq \text{const} \, \| \phi \|_{L^{p'}} \| \psi \|_{L^p}$$

for $|\alpha| \leqq t$ and $|\beta| \leqq -s$ and ϕ, ψ in C^∞ with compact support. This is exploited in [12], and in the discussion of coordinate changes in [11].

Theorem 3 sums up the fundamental connections between A and $\sigma(A)$, given in [8], [9], [11] and [12].

THEOREM 3. *Let A and B be CZO's. Then*

i) *$A + B$ is a CZO and $\sigma(A+B) = \sigma(A) + \sigma(B)$.*

ii) *AB is a CZO and*

$$\sigma(AB)(x, \zeta) = \sum_\alpha [D_\zeta^\alpha \sigma(A)(x, \zeta)](iD_x)^\alpha \sigma(B)(x, \zeta)/\alpha!$$

with the terms arranged according to degree of homogeneity.

iii) *If A^* is defined by $\int (Af) g^* = \int f (A^* g)^*$, where g^* is the complex conjugate of g, then A^* is a CZO and $\sigma(A^*)(x, \zeta) = \sum_\alpha D_x^\alpha (iD_\zeta)^\alpha \sigma(A)(x, \zeta)^*/\alpha!$, arranged by degree of homogeneity.*

iv) *$\sigma(A) = 0$ if and only if A is of order $-\infty$.*

v) *If a_j is any sequence of homogeneous symbols, a_j of degree j and $a_j = 0$ for j sufficiently positive, then there is a CZO, A, with $\sigma(A) = \sum a_j$.*

From (iii) and (iv) we obtain the following. If $\phi\psi = 0$ for some ϕ and ψ in $\mathcal{S}$, and if M_ϕ denotes the map $f \to \phi f$, then $\sigma(M_\phi A M_\psi) = 0$ for any CZO A, so $M_\phi A M_\psi$ is of order $-\infty$. This is described by saying that A is "semi-local." One may further show that an operator B of order $-\infty$ has the form $Bf(x) = \int f(y) b(x, y) \, dy$ for a C^∞ function b. Thus the kernel (in the sense of L. Schwartz) of A is C^∞ "off the diagonal."

Before putting these operators on manifolds, we need a statement about coordinate changes. Let ψ be a diffeomorphism of an open set V in R^n into R^n, let ϕ_j $(j = 1, 2)$ have compact support in V; and define M_ϕ by $M_\phi f = \phi f$, ψ_* by $(\psi_* f)(x) = f(\psi^{-1}(x))$, ψ^* by $(\psi^* f)(x) = f(\psi(x))$.

THEOREM 4. *If A is a CZO of order $\leqq t$, then so is $B = M_{\phi_1} \psi^* A \psi_* M_{\phi_2}$,* *and* $\sigma_t(B)(x, J(x)\zeta) = \phi_1(x)\phi_2(x)\sigma_t(A)(\psi(x), \zeta)$, *where* $J(x)$ *is the transpose of the Jacobian* $\partial\psi/\partial x$.

The proof given in [12] applies here, in view of the remarks after Definition 4.

Now let X be a C^∞ compact manifold without boundary, and $T^*(X)$ be the cotangent bundle with the origin deleted from each fibre. A linear functional f on $C^\infty(X)$ is in $L_s^p(X)$ if for each ϕ in $C^\infty(X)$ with support in a coordinate neighborhood V with coordinates $\chi\colon V \to R^n$, the map $\psi \to f(\phi \cdot (\psi \circ \chi))$ is in $L_s^p(R^n)$, i.e. $|f(\phi \cdot (\psi \circ \chi))| \leqq C \| \psi \|_{L_{-s}^{p'}}$ for ψ in $\mathcal{S}(R^n)$. (If X is not compact, this space is $L_s^p(X, \text{loc})$.) A like definition is given for $B_s^p(X)$. We define, for a coordinate map χ, $\chi^*\psi = \psi \circ \chi$ and $\chi_*\phi = \phi \circ \chi^{-1}$.

Definition 5. A linear map A on $C^\infty(X)$ is a CZO if and only if, when ϕ and ψ have support in a coordinate neighborhood with coordinates χ, then $\chi_* M_\phi A M_\psi \chi^*$ is a CZO on $\mathcal{S}(R^n)$.

If the order of all such $\chi_* M_\phi A M_\psi \chi^*$ is $\leqq t$, then the order of A is $\leqq t$, and $\sigma_t(A)$ is the complex function on $T^*(X)$ defined as follows: if $\zeta_x = \sum z_j d\chi_j(x)$ and $z = (z_1, \cdots, z_n)$, then

$$\phi(x)\psi(x)\sigma_t(A)(\zeta_x) = \sigma_t(\chi_* M_\phi A M_\psi \chi^*)(\chi(x), z).$$

(This differs from the definition in [11], where σ_t is $(-i)^t$ times the definition given here.)

Remarks. It follows from Theorem 3(iii) that for a given χ and a point x where $\phi(x)\psi(x) \neq 0$, $\phi(x)^{-1}\psi^{-1}\sigma_t(\chi_* M_\phi A M_\psi \chi^*)(\chi(x), z)$ is independent of ϕ and ψ; and from Thtorem 4 it follows that if χ' is a different coordinate system and $\sum z_j d\chi_j(x) = \sum z_j' d\chi_j'(x)$, then

$$\sigma_t(\chi'_* M_\phi A M_\psi \chi'^*)(\chi'(x), z') = \sigma_t(\chi_* M_\phi A M_\psi \chi^*)(\chi(x), z).$$

By allowing disconnected coordinate patches, we may have a partition of unity $\sum \phi_j = 1$ with the support of each pair ϕ_j, ϕ_k contained in a single coordinate neighborhood. Thus definition 5 gives a complete description of $A = \sum_{j,k} M_{\phi_j} A M_{\phi_k}$. In particular, $M_\phi A M_\psi$ is of order $-\infty$ when $\phi\psi = 0$, ϕ, ψ in C^∞.

By localizing, Theorem 3 may be transferred to X in the weaker form: $\sigma_t(A)\sigma_s(B) = \sigma_{t+s}(AB)$, $\sigma_t(A) + \sigma_t(B) = \sigma_t(A + B)$, $\sigma_t(A^*) = \sigma_t(A)^*$ with A^* defined by $\int_X (Au)v^* = \int_X u(A^*v)^*$ and $\int_X$ taken with respect to any

C^∞ volume element. More details are given in [12]. The general approach is that of the proof of Lemma 1 below.

LEMMA 1. *Suppose A is of order $\leqq t$ and $|\sigma_t(A)| > 0$ on $T^*(X)$. Then there is a CZO, B, of order $-t$, with $BA - I$ and $AB - I$ of order $-\infty$. If ϕ and ψ have support in a neighborhood V with coordinates χ, and ϕ and $\psi = 1$ on the neighborhood $U \subset V$, then for x in U, we have*

$$\sigma(\chi_* M_\phi B M_\psi \chi^*)(\chi(x), z) = \sum_{j \leqq -t} b_j(y, z)$$

where $y = \chi(x)$, the b_j satisfy

$$\sum_{j+k-|\alpha|=m} (D_z{}^\alpha b_j(y,z))(iD_y)^\alpha a_k(y,z)/\alpha! = \delta_{0m}, \tag{5}$$

and $\sigma(\chi_* M_\phi A M_\psi \chi^*)(y,z) = \sum_{k \leqq t} a_k(y,z)$.

Proof. Let V be a neighborhood in X with coordinates χ, let ϕ, ψ and θ be C^∞ functions supported in V with $\phi\psi = \phi$ and $\psi\theta = \psi$. Let $\sigma(\chi_* M_\theta A M_\theta \chi^*) = \sum_{j \leqq t} a_j$, and define the b_j by (5); for example, $b_{-t} = 1/a_t$, and $b_{-t-1}a_t$ is then determined by (5) with $m = -1$, etc. Then by Theorem 3(v) there is a C with $\sigma(C) = \sum b_j$ on the support of $\chi_*\phi$ and $\chi_*\psi$. Set $B = M_\psi \chi^* C \chi_* M_\psi$. Since $M_\psi A M_{1-\theta}$ is of order $-\infty$, the same is true of $BAM_\theta - BA$, so $\sigma(\chi_* M_\phi BAM_\theta \chi^*) = \chi_*\phi = \sigma(\chi_* M_\phi \chi^*)$, and $M_\phi BAM_\theta - M_\phi$ is of order $-\infty$. Now take V_j to be a covering of X, $\sum \phi_j = 1$, ϕ_j supported in V_j, and define ψ_j, θ_j, B_j as above. Setting $B = \sum M_{\phi_j} B_j$, we find that $BA - \sum M_{\phi_j} B_j A M_{\theta_j}$ and $M_{\phi_j} B_j A M_{\theta_j} - M_{\phi_j}$ are of order $-\infty$. Since $I = \sum M_{\phi_j}$, the lemma is proved.

From Lemma 1 follow the regularity theorems and *a priori estimates*: if f is any distribution on X, and A is elliptic of order t, then $\| f \|_{L_{s+t}{}^p(X)} \leqq C(\| Af \|_{L_s{}^p(X)} + \| f \|_{L_k{}^p(X)})$ for a constant $C < \infty$ depending on A, s, and k, but not on f. This follows immediately from $f = BAf + (I - BA)f$, with B of order $-t$ and $I - BA$ of order $-\infty$.

4. Statement of main theorem (special case). Denote by G_+ a compact manifold with C^∞ boundary $X \subset G_+$. Let G be the double of G_+, with G_- the reflection of G_+ in X, so $G = G_+ \cup G_-$, $G_+ \cap G_- = X$. The spaces $L_s{}^p(G_+)$ are the restrictions to $\operatorname{int}(G_+)$ of distributions in $L_s{}^p(G)$.

Near X, G is of the form $X \times (-1, 1)$; we suppose a normal coordinate t is chosen once and for all, and near X we denote points of G by (x, t), x in X, $-1 < t < 1$. Here $t \geqq 0$ in G_+ and $t \leqq 0$ in G_-, $t = 0$ only on X. Then define R_t by

$$(R_t f)(x) = (f(x,t), (D_t f)(x,t), \cdots, (D_t^{m-1} f)(x,t)),$$

where $D_t = -i\partial/\partial t$. R_t maps $L_s^p(G)$ continuously into $\overset{m-1}{\underset{j=0}{\oplus}} B_{s-j-1/p}^p(X)$, for $s > m - 1 + 1/p$.

In § 4-§ 7, A denotes an invertible elliptic CZO of order m such that, for functions with support near X, A is a differential operator. From Lemma 1 it follows that A^{-1} is a CZO of order $-m$; for there is a B such that $BA - I$ is of order $-\infty$, and multiplying on the right by A^{-1} shows that $B - A^{-1}$ is of order $-\infty$. It follows, further, that $\sigma(B) = \sigma(A^{-1})$.

Denote by $K(A, s, p)$ the space of distributions u in $L_s^p(G_+)$ such that $Au = 0$ in the interior of G_+, and by $H(A, s, p)$ the closure in $\overset{m-1}{\underset{0}{\oplus}} B_{s-1/p-j}^p(X)$ of the space of functions $R_0 u$ for u in C^∞ and $Au = 0$ in G_+. $H(A, s, p)$ resembles the Hardy class H^p, where A is $\partial/\partial\bar{z}$ and the boundary values are taken in a different space.

The main results of the present paper are the following, which extend the announcement in [6].

THEOREM 5. *There is a continuous linear map* $P\colon \overset{m-1}{\underset{0}{\oplus}} B_{s-1/p-j}^p(X) \to K(A, s, p)$. *For each* f, Pf *is* C^∞ *in the interior of* G_+, *and* $\lim_{t\to 0+} R_t P f = P^+ f$ *exists in norm in* $\overset{m-1}{\underset{0}{\oplus}} B_{s-1/p-j}^p(X)$. P^+ *is an* $m \times m$ *matrix of CZO's, and is a projection of* $\oplus B_{s-1/p-j}^p(G_+)$ *onto* $H(A, s, p)$.

THEOREM 6. *If* u *is in* $K(A, s, p)$ *for any* s *real, then* $\lim_{t\to 0+} R_t u$ *exists in norm in* $\overset{m-1}{\underset{0}{\oplus}} B_{s-1/p-j}^p(X)$, *and* $H(A, s, p)$ *is the space of such limits.*

In Section 8 we show how to extend the result to the case that A acts on sections of a bundle, is not assumed invertible, and is defined only in G_+.

5. Construction of P and P^+, and basic lemmas. We suppose, as in the previous section, that a fixed normal variable t has been chosen. It is convenient to assume that t is defined throughout G, $t \geqq 0$ in G_+, $t \leqq 0$ in G_-, although t is a coordinate only near X. We further assume a fixed C^∞ volume elements v on X, and a C^∞ volume element on G which as the form $\int_G f = \int_X \int f(x,t)\,dt\,dv$ for functions f supported near X. As before, A is an invertible elliptic operator on G, of order m, and A^* is defined by

$\int_G (Af)g^* = \int_G f(A^*g)^*$, with g^* the complex conjugate of g. Near X, A has the form

$$Af(x,t) = \sum_0^m A_j(t) D_t^{m-j} f(x,t) \tag{6}$$

where for each t, $A_j(t)$ is a differential operator of order j on X, and $D_t = -i\partial/\partial t$. For $|\epsilon| < 1$ we let G_ϵ be the manifold where $t \geqq \epsilon$. Then, integrating by parts, we find for smooth u and v

$$\int_{G_\epsilon} u(A^*v)^* = \int_{G_\epsilon} (Au)v^* + \int_X (\mathfrak{A}_\epsilon R_\epsilon u, R_\epsilon v), \tag{7}$$

where $(f,g) = \sum_1^m f_j g_j^*$, $g^* =$ complex conjugate of g, and $\mathfrak{A}_\epsilon$ is a matrix of the form

$$\mathfrak{A}_\epsilon = -i \begin{bmatrix} A_{m-1,1}(\epsilon), \cdots\cdots\cdots, A_0(\epsilon) \\ A_{m-2,2}(\epsilon), \cdots, A_0(\epsilon), \quad 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ A_0(\epsilon), \quad 0 \cdots\cdots, \quad 0 \end{bmatrix} \tag{7a}$$

Here A_0 is as in (6), and $A_{j,k} = A_j$ plus terms of order less than j, A_j as in (6). (In the integration by parts, the boundary terms obtain a minus sign because they occur at the lower limit of the $\int dt$.) Since A is elliptic, $A_0(\epsilon)$ is a non-vanishing function and the operator $\mathfrak{A}_\epsilon$ is invertible, the inverse being a matrix of differential operators.

R_ϵ^*, the adjoint of R_ϵ, is continuous from $\bigoplus_0^{m-1} B_{j-1/p-s+1}{}^p(X) \to L_{-s}{}^p(G)$ $(s > m - 1/p)$. We set

$$Pf(x) = (A^{-1}R_0^*\mathfrak{A}_0 f)(x), \qquad x \in \text{interior of } G_+. \tag{8}$$

Thus P is a multiple layer potential with m "layers." In case A is the Laplacian, P is the familiar combination of single and double layer potentials whereby a harmonic function is represented in terms of its Cauchy data on the boundary.

The essential part of the proof of Theorem 5 is Lemma 2 below. It resembles some of the computations that arise in various derivations of *a priori* estimates and definitions of ellipticity, as in Browder [2], [3]. Here, however, the problem is not reduced to constant coefficients, and boundary conditions are not imposed.

Define $R_{1,\epsilon}$ by $(R_{1,\epsilon}f)(x) = f(x,\epsilon)$.

LEMMA 2. *If g is in $C^\infty(X)$, then for each j and k*

$$R_{1,\epsilon}D_t{}^kA^{-1}D_t{}^jR_{1,0}{}^*g \tag{9}$$

converges uniformly to a limit $L_{kj}g$, *as* $\epsilon \to 0+$. *Further, all derivatives of* (9) *converge uniformly.* L_{kj} *is a CZO of order* $\leqq j+k-m+1$, *and*

$$\sigma_{j+k-m+1}(L_{kj}) = \tfrac{1}{2\pi}\int_\Gamma \frac{\tau^{j+k}}{\sum\limits_0^m \sigma_l(A_l(0))\tau^{m-l}}\,d\tau \tag{10}$$

with A_l *as in* (6), *and* Γ *a path surrounding those roots of the polynomial* $\sum\limits_0^m \sigma_l(A_l(0))\tau^{m-l}$ *lying in* $\{\mathrm{Im}(\tau) > 0\}$.

Proof. It suffices to consider $M_{\phi_1}A^{-1}M_{\phi_2}R_{1,0}{}^*g$ when either the supports of ϕ_1 and ϕ_2 lie in $\{|t| < 1\}$, or when their supports are disjoint, or when the support of ϕ_2 does not intersect X. The last case is trivial, since $M_{\phi_2}R_{1,0}{}^*g = 0$. When ϕ_1 and ϕ_2 have disjoint supports, $D_t{}^kM_{\phi_1}A^{-1}M_{\phi_2}D_t{}^jR_{1,0}{}^*$ is continuous from $B_{s-1/p}{}^p$ to $L_l{}^p$ for every l and for $s < 1/p$, since $M_{\phi_1}A^{-1}M_{\phi_2}$ is of order $-\infty$. Taking $p = 2$ and $l > 2/n$ we have from a well-known theorem of Soboleff (or from an extreme case of Theorem 1) that $D_t{}^kM_{\phi_1}A^{-1}M_{\phi_2}D_t{}^jR_{1,0}{}^*g$ is continuous, so $R_{1,\epsilon}D_t{}^kM_{\phi_1}A^{-1}M_{\phi_2}D_t{}^jR_{1,0}{}^*g$ converges uniformly. From the restriction theorem, the limit $R_{1,0}D_t{}^kM_{\phi_1}A^{-1}M_{\phi_2}D_t{}^jR_{1,0}{}^*$ is of order $-\infty$.

Suppose now ϕ_1 and ϕ_2 have support in a neighborhood $U \subset G$ with coordinates $\chi\colon U \to R^n$, such that t is one of the coordinates, i.e. $\chi = (t, \chi_2, \cdots, \chi_n)$. Let $\phi = 1$ on the support of $|\phi_1| + |\phi_2|$, and support$(\phi) \subset U$. From Lemma 1 and the remark that the operator B of that lemma differs from A^{-1} by an operator of order $-\infty$, we deduce the form of A^{-1} in local coordinates. Specifically, let

$$C = \chi_*M_\phi A^{-1}M_\phi\chi^*,$$

and

$$\sigma(\chi_*M_\phi AM_\phi\chi_*)\,(t, y; \tau, \eta) = \sum_0^m a_j(t, y; \tau, \eta)$$

with a_j homogeneous of degree j in (τ, η). Here $\eta = (\eta_2, \cdots, \eta_n)$ are the variables "dual to" y, and τ is "dual to" t. Then on χ(support$(|\phi_1| + |\phi_2|)$) we have $\sigma(C) = \sum\limits_{l \leqq -m} c_l$ where the c_l satisfy

$$c_{-m}a_m = 1$$

$$\begin{aligned} c_{-m}a_{m-1} + c_{-m-1}a_m - i(\partial c_{-m}/\partial\tau)\,(\partial a_m/\partial t) \\ - i\sum_2^n (\partial c_{-m}/\partial\eta_j)\,(\partial a_m/\partial y_j) = 0 \text{ etc.,} \end{aligned} \tag{11}$$

from (5).

Now set

$$C_l = M_{\phi_1}\chi^* \mathrm{Op}(c_l, \theta)\chi_* M_{\phi_2},$$

where $\theta(\eta, \tau) = 1$ if $|\eta|^2 + |\tau|^2 \geqq 1$, and θ vanishes near $\eta = 0$, $\tau = 0$. Then if D is any operator of order $\leqq q$, $D[M_{\phi_1}A^{-1}M_{\phi_2} - \sum_{l > l_0} C_l]D_t^j$ has order $\leqq l_0 + q + j$ so that for $l_0 \leqq j - 1 - q - \frac{n+1}{2}$ we obtain again from Soboleff's theorem the continuity of $D[M_{\phi_1}A^{-1}M_{\phi_2} - \sum_{l > l_0} C_l]D_t^j R_{1,0}{}^* g$, and hence the fact that it converges uniformly to its boundary values. Thus we are reduced to the analysis of $R_{1,\epsilon}D_t^k C_l D_t^j R_{1,0}{}^*$.

Let $\alpha \in C^\infty(R^1)$, $\alpha(s)$ vanish except for $-1 < s < -\frac{1}{2}$, α real, and $\int \alpha = 1$. Set $\alpha_m(s) = m\alpha(ms)$. Then as a distribution, $\alpha_m \otimes g \rightarrow R_{1,0}{}^* g$, i.e. $\int\int \alpha_m(t) g^*(y)\phi(t, y)\, dy dt \rightarrow \int g^*(y)\phi(0, y)\, dy$ for ϕ continuous. It follows that $m^{j+1}\alpha^{(j)}(mt)g(y) \rightarrow D_t^j R_{1,0}{}^* g$, when $\alpha^{(j)} = D_t^j \alpha$. Thus with

$$f = \chi_* \phi_2 g \text{ and } \phi_{1*} = \chi_* \phi_1$$

we have

$$(12) \qquad (\chi_* C_l D_t^j R_{1,0}{}^* g)(t, y)$$

$$= (2\pi)^{-n} \phi_{1*} \lim_{m \to \infty} \int_{R^{n-1}} \int_{-\infty}^{\infty} \theta \tau^j \alpha^\wedge (\tau/m) f^\wedge c_l e^{it\tau} d\tau e^{i\langle y, \eta\rangle} d\eta$$

on every open set where this expression converges uniformly. We obtain from (12) an integral in which $\lim_{m \to \infty}$ may be taken under the integral sign by replacing $\int_{-\infty}^{\infty} d\tau$ with a contour integral over a finite path in $\{\mathrm{Im}(\tau) \geqq 0\}$. The necessary holomorphic extension is obtained from (11), where $a_m(t, y; \tau, \eta)$ is a polynomial in τ of degree m, and non-vanishing for $|\tau|^2 + |\eta|^2 > 0$ by definition of ellipticity of A. Allowing now τ to be complex, there is a compact set C in the complex plane $\mathbf{C}$, not intersecting the real axis, such that for $|\eta| = 1$ and (t, y) in support$(\chi(|\phi_1| + |\phi_2|))$ the zeroes of a_m lie in C. Thus $c_{-m} = 1/a_m$ has a holomorphic extension to the complement of C, for such η and (t, y). Then it follows from (11) by induction that all the c_l and all their derivatives have such extensions, each of which is bounded by some polynomial in $|\tau|$. Suppose C lies in $\{|\tau| \leqq r\}$, and set

$$(12a) \qquad \Gamma(\eta) = \text{boundary of } \{|\tau| \leqq \max(1, r|\eta|)\} \cap \{\mathrm{Im}(\tau) \geqq 0\}.$$

Extend θ to $R^{n-1} \times \mathbf{C}$ so that $\theta(\eta, \tau)$ still $= 1$ for $|\eta|^2 + |\tau|^2 \geqq 1$. Since $\alpha^\wedge(\tau) = \int_{-1}^{-\frac{1}{2}} e^{-it\tau}\alpha(t)\, dt$ decays exponentially as $|\tau| \rightarrow \infty$ in $\{\mathrm{Im}(\tau) \geqq 0\}$, and $e^{i\tau t}$ is bounded for $t \geqq 0$ and $\mathrm{Im}(\tau) \geqq 0$, and the integrand in (12) is holomorphic between $\Gamma(\eta)$ and $\{\mathrm{Im}(\tau) = 0\}$ and bounded by a polynomial in $|\tau|$, we may replace $\int_{-\infty}^{\infty} d\tau$ by $\int_{\Gamma(\eta)} d\tau$. Then $\lim_{m\to\infty}$ may pass inside the integrals, yielding for (12) the expression

$$(2\pi)^{-n}\phi_{1*}\int_{R^{n-1}}\int_{\Gamma(\eta)}\theta c_l\tau^j e^{it\tau}d\tau f^\wedge e^{i\langle y,\eta\rangle}d\eta, \tag{13}$$

since $\alpha^\wedge(\tau/m)$ converges boundedly to $\alpha^\wedge(0) = \int\alpha = 1$ as $m \to \infty$. It is clear that (13) is C^∞ for $t > 0$, and all its derivatives in y and t extend continuously to $\{t \geqq 0\}$. Thus the proof of the lemma is completed by showing that $\lim_{\epsilon\to 0+} R_{1,\epsilon}D_t^kC_lD_t^jR_{1,0}^*$ is a CZO of order $\leqq l + k + j + 1$, and computing its symbol when $l = -m$.

From (13) we have

$$\begin{aligned}(14)\qquad &(\lim_{\epsilon\to 0+} R_{1,\epsilon}D_t^kC_lD_t^jR_{1,0}^*g)(x)\\ &= (2\pi)^{1-n}\phi_{1*}\int_{R^{n-1}}\Big[\tfrac{1}{2\pi}\int_{\Gamma(\eta)}\theta c_l\tau^{j+k}d\tau\Big]f^\wedge e^{i\langle y,\eta\rangle}d\eta\\ &\qquad + \text{similar terms involving lower powers of } \tau \text{ and derivatives}\\ &\qquad\quad \text{of } c_l \text{ with respect to } t;\ \text{here } y = \chi(x).\end{aligned}$$

Now the function in square brackets here is homogeneous of degree $l+j+k+1$ when $|\eta| \geqq \max(1, 1/r)$; for the homogeneity of c_l is preserved in its holomorphic extension, and with the given values of η, $\theta(\tau,\eta) = 1$ and $\Gamma(\eta)$ is the boundary of $\{|\tau| \leqq r|\eta|\} \cap \{\mathrm{Im}(\tau) > 0\}$. This shows that (14) is a CZO of order $\leqq l + j + k + 1$, since the required differentiability is clear. Finally, taking $l = -m$ and setting $c_{-m} = 1/a_m$ in (14), we have

$$\begin{aligned}&\sigma_{-m+j+k+1}(\chi_*\lim R_{1,\epsilon}D_t^kM_{\phi_1}A^{-1}M_{\phi_2}D_t^jR_{1,0}^*\chi^*)(y,\eta)\\ &\qquad = \frac{1}{2\pi}\int_{\Gamma(\eta)}\frac{\tau^{j+k}}{a_m(0,y;\tau,\eta)}\,d\tau\,\phi_1(\chi^{-1}(0,y))\phi_2(\chi^{-1}(0,y)).\end{aligned}$$

q. e. d.

LEMMA 3. *If* $f \in B_{s+j+k+1-m}{}^p$ *then* $R_{1,\epsilon}D_t^kA^{-1}D_t^jR_{1,0}^*f \to L_{kj}f$ *in* $B_s{}^p$.

Proof. It is enough to show that

$$\|R_{1,\epsilon}D_t^kA^{-1}D_t^jR_{1,0}^*f : B_s{}^p\| \leqq c\,\|f : B_{s+j+k+1-m}{}^p\|$$

for f in C^∞, with c independent of ϵ. For then the same uniform bound holds with f in $B_{s+j+k+1-m}{}^p$, and Lemma 2 shows that the desired convergence holds for f in the dense subset of C^∞ functions, so the lemma follows from a standard argument.

The proof of Lemma 2 shows that it is sufficient to obtain a bound, independent of t for $|t| < 1$, on the homogeneous terms

$$\text{(i)} \qquad \phi_* \int_{R^{n-1}} c_{ljk} e^{i\langle y,\eta\rangle} f^\wedge(\eta)\, d\eta \qquad (l \leqq -m)$$

with

$$c_{ljk}(t, y; \tau, \eta) = \int_{\Gamma(\eta)} \theta(\tau, \eta) \tau^{j+k} c_l(t, y; \tau, \eta) e^{it\tau} d\tau$$

as in (14). Assume that $\phi_*(t, y)$ vanishes for $|\delta y| > \frac{1}{2}$, and let $e_\alpha(y) = e^{\pi i \langle \alpha, y\rangle \delta}$. Choose a function ψ supported in $|\delta y| \leqq 1$ with $\psi\phi_* = \phi_*$. Then

$$\phi_*(t, y) c_{ljk}(t, y; \eta) = \phi_*(t, y) \sum b_\alpha(t, \eta) e_\alpha(y)$$

with $b_\alpha = (2\delta)^{n-1} \int_{R^{n-1}} \psi c_{ljk} e_\alpha{}^*$. According to Corollary 2.1, the multiplication operator M_{e_α} on $B_s{}^p(R^{n-1})$ has a norm which is dominated by $\max |D^\beta e_\alpha|$ for $|\beta| \leqq |s| + 1$; thus $\| M_{e_\alpha} \| = O(1 + |\alpha|^{|s|+1})$. Now if B_α is defined by $(B_\alpha f)^\wedge(\eta) = b_\alpha(t, \eta) f^\wedge(\eta)$, we have that (i) is $\phi_* \sum M_{e_\alpha} B_\alpha f$; thus it remains only to find appropriate bounds on the norm of B_α as an operator from $B_{s+j+k+1-m}{}^p$ to $B_s{}^p$. According to Mihlin's theorem, and the fact that J^t is an isomorphism of $B_{s-t}{}^p$ with $B_s{}^p$, this norm is dominated by the supremum for η in R^{n-1} and $|\beta| \leqq n$ of

$$\text{(ii)} \qquad |\eta|^{|\beta|} [(1 + |\eta|^2)^{m-j-k-1} b_\alpha(t, \eta)]|.$$

For $|\eta| \leqq r$ (with r as in (12a)) we have $\Gamma(\eta) = \Gamma(r)$ so, integrating by parts, we find for $|\eta| \leqq r$ that

$$\text{(iii)} \qquad \text{(ii)} \leqq C_\nu (1 + |\alpha|^2)^{-\nu} \text{ for each integer } \nu > 0;$$

here C_ν depends on the supremum of finitely many derivatives of $\psi\theta c_l$ in $|\eta| \leqq r$. For $|\eta| \geqq r$,

$$\int_{\Gamma(\eta)} \theta \tau^{j+k} c_l e^{it\tau}\, d\tau$$
$$= (|\eta|/r)^{j+k+l+1} \int_{\Gamma(\eta)} \sigma^{j+k} c_l(t, y; r\eta/|\eta|, \sigma) e^{it\sigma|\eta|/r}\, d\sigma;$$

since $l \leqq -m$, it follows that an estimate of the form (iii) holds also for $|\eta| \geqq r$. Choosing ν so that $\sum_\alpha (1 + |\alpha|)^{|s|+1-\nu} < \infty$ completes the proof.

Let now

$$\| f \|_s = \sum_0^{m-1} \| f_j \|_{B_{s-1/p-j}{}^p}. \tag{15}$$

LEMMA 4. *For $s < 1/p$ there is a constant $C_{s,p}$ such that*

$$\| A^{-1} R_0{}^* \mathfrak{A}_0 f \colon L_s{}^p(G) \| \leqq C_{s,p} \| f \|_s$$

for f in C^∞. For $s \geqq 0$ there is a corresponding bound on the norm in $L_s{}^p(G_+)$ of $Pf = (A^{-1} R_0{}^* \mathfrak{A}_0 f) \mid \operatorname{int}(G_+)$.

Proof. For $s < 1/p$ we have the lemma from bounds on $\mathfrak{A}_0$, R_0, and A^{-1}. We conclude the proof by establishing the estimates for $s = \text{integer} \geqq m$ and interpolating.

For k an integer $\geqq m$, and $0 \leqq j \leqq m-1$, let $g_j = \lim_{t \to 0+} D_t{}^j A^{-1} R_0{}^* \mathfrak{A}_0 f$. This exists uniformly, and $\| g_j \colon B_{k-1/p-j}{}^p \| \leqq C \| f \|_k$, by Lemma 2. Now let g_m be given by $A_0(0) g_m + \sum_1^m A_j(0) g_{m-j} = 0$, g_{m+1} by

$$A_0 g_{m+1} + (A_0{}' + A_1) g_m + \cdots + (A_{m-1}{}' + A_m) g_1 + A_m{}' g_0 = 0$$

(with $A_j{}'$ obtained from the A_j of (6) by applying D_t to its coefficients), and in general let g_j $(m \leqq j \leqq k)$ be the solution for $D_t{}^j w$ in the equations obtained by expanding $D_t{}^r \sum_0^m A_l D_t{}^{m-l} w = D_t{}^r A w = 0$, for $0 \leqq r < k-m$. Then again $\| g_j \colon B_{k-1/p-j}{}^p \| \leqq C \| f \|_k$, so by Theorem 1 we can choose v with $\| v \colon L_k{}^p(G) \| \leqq C \sum_0^{k-1} \| g_j \|$, v in C^∞ for $t \neq 0$, and $R_{1,0} D_t{}^j v = g_j$. Let $u = A^{-1} R_0{}^* \mathfrak{A}_0 f$ in $\{t \geqq 0\}$, and $u = v$ in $\{t < 0\}$. Then $u \in L^p(G)$, so for ϕ in C^∞ we have

$$\int_G u(A^*\phi)^* = \lim_{\delta \to 0} \int_{|t|>\delta} u(A^*\phi)^*$$

$$= \lim \int_{|t|>\delta} (Au)\phi^* + \lim \int_X (R_\delta u, \mathfrak{A}_\delta{}^* R_\delta \phi) - \lim \int_X (R_{-\delta} u, \mathfrak{A}_{-\delta}{}^* R_{-\delta} \phi).$$

Here the two integrals over X cancel as $\delta \to 0$, since v has the same Cauchy data as $A^{-1} R_0{}^* \mathfrak{A}_0 f$ at $t = 0$; this yields $\int_G u(A^*\phi)^* = \int_{t \leqq 0} (Av)\phi^*$, since $Au = 0$ for $t > 0$. Now $Av \in L_{k-m}{}^p(G)$, and by the choice of the g_j, the Cauchy data of order $< k-m$ on $\{t = 0\}$ of Av are zero. It follows from Corollary 1.1 that $Au \in L_{k-m}{}^p(G)$, hence by the regularity theorem (end of Section 3) that $u \in L_k{}^p(G)$, and in fact $\| u \colon L_k{}^p(G) \| \leqq \text{const} \| f \|_k$. Thus the lemma is established for $s = \text{integer} \geqq m$, as well as for $s < 1/p$.

Finally, for an arbitrary $s \geqq 0$, choose an integer $k \geqq \max(m,s)$. Let Ext be a continuous extension map from $L_k{}^p(G_+)$ to $L_k{}^p(G)$; it follows from the proof of Corollary 1.2 that there is such an Ext which is also continuous from $L_0{}^p(G_+)$ to $L_0{}^p(G)$. By what we have shown, $f \to \operatorname{Ext}((A^{-1}R_0{}^*\mathcal{A}_0 f) \mid G_+)$ is continuous from $\oplus B_{t-1/p-j}{}^p$ to $L_t{}^p(G)$ for $t = 0$ and $t = k$. Then by Theorem 2 on interpolation, it is also continuous for $t = s$, and Lemma 3 is proved.

On the basis of Lemma 3, we may define for an m-tuple f of distributions on X,

$$P^{\pm} f = \pm \lim_{\epsilon \to 0\pm} R_\epsilon A^{-1} R_0{}^* \mathcal{A}_0 f \tag{16}$$

Thus $P^+ f = \lim R_\epsilon P f$, with P as in (8). The principal steps remaining for the proof of Theorem 5 are those showing that P^+ and P^- are complementary projections. These generalize classical "jump relations" for harmonic and holomorphic potentials.

LEMMA 5. *$P^+ + P^- = I$.*

Proof. Let $\phi \in C^\infty(G)$, with support near X, and let $\langle \phi, g \rangle$ denote the value of the distribution g on the function ϕ. (This form is taken conjugate linear in g.) We have for an m-tuple f in $C^\infty(X)^m$ that

$$\begin{aligned}\int_X (\mathcal{A}_0 f, R_0\phi) &= \langle R_0{}^*\mathcal{A}_0 f, \phi\rangle = \langle APf, \phi\rangle = \langle Pf, A^*\phi\rangle \\ &= \lim_{\epsilon\to 0+} \int_{|t|>\epsilon} (Pf)\,(A^*\phi)^* = \lim_{\epsilon \to 0+} \int_X (\mathcal{A}_\epsilon R_\epsilon Pf, R_\epsilon \phi) \\ &\quad - \lim_{\epsilon\to 0+} \int_X (\mathcal{A}_{-\epsilon} R_{-\epsilon} Pf, R_{-\epsilon}\phi) = \int_X (\mathcal{A}_0 P^+ f, R_0\phi) + \int_X (\mathcal{A}_0 P^- f, R_0 \phi).\end{aligned}$$

Here the introduction of $\lim\limits_{\epsilon \to 0+}$ is justified by the fact that $Pf \in L_0{}^p(G)$ (Lemma 4), and the last two equalities depend on Lemma 2. Since every C^∞ m-tuple on X has the form $R_0\phi$ for a ϕ in C^∞, and $\mathcal{A}_0$ is invertible, we have $P^+ + P^- = I$, q. e. d.

Now let $H^{\pm} = \{R_0 u \colon u \in L_m{}^p(G_\pm) \text{ and } Au = 0 \text{ in } G_\pm\}$.

LEMMA 6. *$\bigoplus\limits_0^{m-1} B_{m-j-1/p}{}^p(X)$ is the direct sum of H^+ and H^-, and $P^{\pm}$ are the corresponding projections.*

Proof. Because $f = P^+ f + P^- f$, the lemma follows from

(i) range of $P^{\pm}$ lies in $H^{\pm}$, and

(ii) $H^{+} \cap H^{-} = 0$.

But (i) follows from Lemma 4: if $f = P^{+}g$, then $u = A^{-1}R_0{}^{*}\mathcal{A}_0 g$ is in $L_m{}^{p}(G_{+})$ and $Au = 0$ in G_{+}, while $R_0 u = P^{+}g = f$ by (16). For (ii), let $Au = 0$ in G_{+}, and $R_0 u = R_0 v$. Then the function w, ($w = u$ in G_{+}, $w = v$ in G_{-}), is in $L_m{}^{p}(G)$ by Corollary 1.1. An appeal to Green's formula as in Lemma 4 shows that $\langle w, A^{*}\phi\rangle = \lim \int_{|t|>\epsilon} w(A^{*}\phi)^{*} = 0$ for ϕ in C^{∞}, hence that $Aw = 0$, so $w = 0$ and $R_0 u = R_0 w = 0$.

6. Proof of Theorem 5, and computation of the symbol of P^{+}. P is defined by (8). Lemma 4 shows that P is continuous from $\bigoplus_{0}^{m-1} B_{s-j-1/p}$ into $K(A, s, p)$. The claim that Pf is C^{∞} in G_{+} follows from the semi-local nature of the CZO A^{-1} (remarked after Theorem 3). The existence in $\bigoplus_{0}^{m-1} B_{s-j-1/p}{}^{p}$ of $\lim_{\epsilon\to 0+} R_\epsilon Pf$ is given in Lemma 3, and Lemma 2 shows $P^{+} = \lim R_\epsilon P$ to be a matrix of CZO's. (The same lemma allows us to describe the symbol of P^{+}, but first we complete the proof of Theorem 5.)

It remains only to identify $P^{+}(\bigoplus_{0}^{m-1} B_{s-j-1/p}{}^{p})$ as $H(A, s, p)$, the closure of the space of boundary values of C^{∞} solutions of $Au = 0$. First, if u is C^{∞} and $Au = 0$ in G_{+}, then $R_0 u$ is in the space H^{+} introduced in Lemma 6, so $R_0 u = P^{+}R_0 u$ by that lemma. Now if u^{j} is a sequence in $C^{\infty}(G)$, $Au^{j} = 0$ in G_{+}, and $R_0 u^{j} \to f$ in $\bigoplus_{0}^{m-1} B_{s-j-1/p}{}^{p}$, then $f = \lim R_0 u^{j} = \lim P^{+}R_0 u^{j} = P^{+}f$; this shows $H(A, s, p) \subset P^{+}(\bigoplus_{0}^{m-1} B_{s-j-1/p}{}^{p})$. Conversely, if $P^{+}f = f$ we may let $f^{j} \to f$, f^{j} in C^{∞}, and set $u^{j} = Pf^{j}$. Then $Au^{j} = 0$ in G_{+} and $R_0 u^{j} = P^{+}f^{j} \to P^{+}f = f$.

Finally, we give a description of the top-order terms of the symbol of P^{+}. P^{+} is an $m \times m$ matrix of operators mapping $\bigoplus_{0}^{m-1} B_{s-j-1/p}{}^{p}$ into itself, thus $P^{+} = (P^{+}{}_{ij})$ with $P^{+}{}_{ij}$ of order $\leqq i - j$. Denote by $\sigma_{(0)}(P^{+})$ the matrix $(\sigma_{i-j}(P^{+}{}_{ij}))$ of top-order symbols. Now if we reduce the operator $A = \sum_{0}^{m} A_j D_t{}^{m-j}$ to a first order system in D_t in the usual way, we obtain the system

$$D_t + A_0^{-1}\begin{bmatrix} 0, & -A_0, & 0, \cdots\cdots, 0 & \\ 0, & 0, & -A_0, & 0, \cdots, 0 \\ \cdot & & & \\ \cdot & & & -A_0 \\ \cdot & & & \\ A_m, & A_{m-1}, & \cdots\cdots\cdots, & A_1 \end{bmatrix}$$

Let η be a cotangent to X. *Then* $\sigma_{(0)}(P^{\pm})(\eta)$ *is the projection on the generalized eigenspace of*

$$-\sigma_0(A_0)(\eta)^{-1}\begin{bmatrix} 0, & -\sigma_0(A_0)(\eta), & 0 \cdots 0 \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ & & \sigma_1(A_1)(\eta) \end{bmatrix} = \sigma_{(0)}(A^{\sim})(\eta)$$

corresponding to eigenvalues lying in $\{\operatorname{Im}(\tau) \gtrless 0\}$. To prove this observe that $\left[\lambda I + \begin{bmatrix} a_m & \cdots\cdots & a_1 \\ & & \cdot \\ & & \cdot \\ & & \cdot \\ & & -1 \\ 0 & -1 & 0 \cdots 0 \end{bmatrix}\right]^{-1} = (b_{jl})$ $(0 \leqq j, l < m)$ with

$$b_{jl} = \sum_{k=0}^{m-1} \lambda^{j+k}\left[\sum \lambda^{m-i} a_i\right]^{-1} a_{m-k-l-1},$$

where $a_0 = 1$ and $a_j = 0$ if $j < 0$. According to Lemma 2 and formula (7a) for $\mathcal{A}_0$, we then obtain

$$\begin{aligned} &\sigma_{j-l}(P^{+}{}_{jl}) \\ &\quad = \frac{1}{2\pi i}\int_{\Gamma} \sum_{k=0}^{m-1} \tau^{j+k}\left[\sum_{i=0}^{m} \tau^{m-i}\sigma_i(A_i)(\eta)\right]^{-1}\sigma_{m-k-l-1}(A_{m-k-l-1})(\eta)\, d\tau \\ &\quad = \frac{1}{2\pi i}\int_{\Gamma} [\lambda I + \sigma_{(0)}(A^{\sim})(\eta)]^{-1}\, d\eta. \end{aligned}$$

The fact that this is the desired projection follows from properties of analytic functions on Banach algebras, or from the Jordan decomposition.

7. Proof of Theorem 6. Here we prove the existence of $\lim_{\varepsilon \to 0+} R_\varepsilon u$ for u in $K(A, s, p)$, and the fact that $H(A, s, p)$ is the space of such limits.

This is a natural adjunct to Theorem 5, but its proof is separated because it uses a number of results beyond those required by Theorem 5.

The main difficulty is Lemma 7 below, which is treated last; first we state the two main lemmas and show how they yield Theorem 6.

LEMMA 7. *If* $u \in K(A,s,p)$, *then* $R_\epsilon u$ *converges in* $\bigoplus_{0}^{m-1} B_{s-1/p-j}{}^{p}$ *as* $\epsilon \to 0+$, *and* $\| \lim R_\epsilon u \| \leqq \text{const} \| u \|$.

LEMMA 8. *If* $u \in K(A,s,p)$ *and* $f = \lim R_\epsilon u$, *then*

$$u = Pf \text{ and } f \in H(A,s,p).$$

Thus to complete Theorem 6 we need only show that

$$H(A,s,p) \subset R_{0+}(K(A,s,p)).$$

Let $f \in H(A,s,p)$, and $f_j \in H(A,\infty,p)$, $f_j \to f$. By Lemma 4,

$$Pf_j \to Pf \in K(A,s,p);$$

from formula (16) $\lim R_\epsilon Pf = P^+f = f$, so Theorem 6 is proved.

Proof of Lemma 8. First, if $s = m$ then by Lemma 6, $f = R_0 u$ lies in H^+ and $f = P^+f$. If we set $v = Pf - u$ in G_+, then $\lim\limits_{\epsilon \to 0+} R_\epsilon v = P^+f - R_0 u = 0$, so we may extend v to an element of $L_m{}^p(G)$ by setting $v = 0$ in G_-. An application of Green's formula (as in Lemma 6) shows $Av = 0$, so $v = 0$. This proves Lemma 7 for $s = m$.

Now let s be arbitrary, and $u \in K(A,s,p)$. Given a point $g \in G_+$, choose $\epsilon > 0$ so that $g \in \{t > 2\epsilon\} = G_{2\epsilon}$. Since u is C^∞ in G_+, we may apply the case $s = m$ proved above to show that $u = A^{-1}R_\epsilon{}^*\mathcal{A}_\epsilon R_\epsilon u$ in G_ϵ. Choose ϕ and ψ in $C^\infty(G)$ so that $\phi\psi = 0$, $\phi(g) = 1$, and $\psi = 1$ in $\{|t| < 2\epsilon\}$. Then $u(g) = (M_\phi A^{-1} M_\psi R_\epsilon{}^* \mathcal{A}_\epsilon R_\epsilon u)(g)$. Since $M_\phi A^{-1} M_\psi$ is of order $-\infty$, $R_\epsilon{}^*\mathcal{A}_\epsilon R_\epsilon u \to R_0 \mathcal{A}_0 f$ in $L_t{}^p$ for $t = \min(s,0)$, and the map $v \to v(g)$ is continuous on $L_t{}^p$ for $t > n/p$, we have

$$u(g) = (M_\phi A^{-1} M_\psi R_\epsilon{}^* \mathcal{A}_\epsilon R_\epsilon u)(g) \to (M_\phi A^{-1} M_\psi R_0{}^* \mathcal{A}_0 f)(g) = (Pf)(g).$$

Then $P^+f = \lim\limits_{\epsilon \to 0+} R_\epsilon Pf = \lim\limits_{\epsilon \to 0+} R_\epsilon u = f$. Finally, let f^j be a sequence of C^∞ m-tuples, $f^j \to f$ and set $u^j = Pf^j$. Then $Au^j = 0$ in G_+, and $R_0 u^j = P^+f^j \to P^+f = f$, so $f \in H(A,s,p)$.

We are left with the proof of Lemma 7, which proceeds as follows. If $u \in K(A,s,p)$ then there is a distribution $v \in L_s{}^p(G)$, agreeing with u in G_+, and consequently $w = Av \in L_{s-m}{}^p(G)$ and vanishes in G_+. Thus we want to show that if $w \in L_{s-m}{}^p(G)$ and w vanishes in the interior of G_+, then $\lim R_\epsilon A^{-1} w$

exists in the appropriate space. The question is reduced to Euclidean space by choosing a resolution of the identity $\sum \phi_j = 1$, and functions ψ_j each supported in a coordinate neighborhood and such that $\psi_j = 1$ on the support of ϕ_j. Since $A^{-1} = \sum \psi_j A^{-1} \phi_j + \sum (1 - \psi_j) A^{-1} \phi_j$, and the second sum is an operator of order $-\infty$, it suffices to show that $Bu = \psi A^{-1} \phi u$ has the desired normal derivatives if u vanishes inside G_+, $\psi = 1$ on the support of ϕ, and ψ is supported in a coordinate neighborhood.

Notice that if Bu has normal derivatives of order $< m$, then it has normal derivatives of all orders; for $Bu = A^{-1}\phi u + (\psi - 1) A^{-1} \phi u$, with $(\psi - 1) A^{-1} \phi u$ in C^∞, and since u vanishes in G_+, $A(A^{-1}\phi u) = 0$ in G_+, which allows us to solve for the normal derivatives of order $\geqq m$ in terms of those of order $< m$.

Without introducing any new notation, we take B to be a CZO on Euclidean space, and u to be in $L_s^p(R^{n+1})$ with support in the half space $\{t \leqq 0\}$. For coordinates in R^{n+1} we take $(x_1, \cdots, x_n, t)$, and in the dual space $(\xi_1, \cdots, \xi_n, \tau)$. Define an operator Λ_+ by

$$(\Lambda_+ f)^\wedge (\xi, \tau) = [(1 + \sum_1^n \xi_j^2)^{1/2} - i\tau] f^\wedge (\xi, \tau). \tag{18}$$

It follows from Mihlin's theorem that $\Lambda_+ J^1$ is an isomorphism of L_s^p onto L_s^p for all s and $1 < p < \infty$, with the J^1 of Definition 1, and hence that Λ_+ is an isomorphism of L_s^p onto L_{s-1}^p.

The point of using Λ_+ instead of J^{-1} is the following.

LEMMA 9. *If u is any distribution in $S'(R^{n+1})$ with support in $\{t \leqq 0\}$, then $(\Lambda_+)^s u$ also has support in $\{t \leqq 0\}$.*

Proof. Let $\langle \phi, u \rangle$ be the duality between $\mathcal{S}$ and $\mathcal{S}'$, taken as conjugate linear in u. Then by Parseval's equality, $\langle \phi, \Lambda_+^s u \rangle = \langle \Lambda_-^s \phi, u \rangle$ with $(\Lambda_-^s \phi)^\wedge (\xi, \tau) = [(1 + \sum_1^n \xi_j^2)^{1/2} + i\tau]^s \phi (\xi, \tau)$. Thus it suffices to show that if $\phi \in \mathcal{S}$ and vanishes for $\{t \leqq \epsilon\}$ for some $\epsilon > 0$, then $\Lambda_-^s \phi$ vanishes for $\{t \leqq 0\}$. But $\phi^\wedge(\xi, \tau) = \int_{R^n} \int_\epsilon^\infty e^{-\langle x, \xi \rangle - it\tau} \phi(x, t)\, dt dx$ has a holomorphic extension to $\{\operatorname{Im} \tau < 0\}$; in fact $\phi^\wedge(\xi, a - ib)$ is the Fourier transform of $e^{-tb}\phi(x, t)$, which $\rightarrow 0$ in $\mathcal{S}$ as $b \rightarrow +\infty$. Thus we have for $b \geqq 0$ that

$$(\Lambda_-^s \phi)(x, t)$$

$$= (2\pi)^{-n-1} \int_{R^n} \int_{-\infty}^{\infty} e^{i\langle x, \xi \rangle + ita} e^{tb} [(1 + \sum_1^n \xi_j^2)^{1/2} + ia + b]^s \phi^\wedge (\xi, a - ib)\, d\tau d\xi,$$

by Cauchy's theorem. Taking $t \leqq 0$ and letting $b \to +\infty$ yields $(\Lambda_-^s\phi)(x,t) = 0$.

Unfortunately Λ_+ is not a CZO, for $\sigma_1(\Lambda_+)$ would be $(\sum_n^1 \xi_j^2)^{1/2} - i\tau$, which is not C^∞ on $\sum \xi_j^2 + \tau^2 = 1$. However, Lemma 10 shows that in some respects it behaves like one. When A is any operator on $\mathcal{S}$, we define its successive commutators with Λ_+ by:

(19) $\qquad C_0(A,\Lambda_+) = A, \quad C_{j+1}(A,\Lambda_+) = C_j(A,\Lambda_+)\Lambda_+ - \Lambda_+ C_j(A,\Lambda_+).$

LEMMA 10. *If A is a CZO of order $\leqq r$, then $C_j(A,\Lambda_+)$ has order $\leqq r$.*

Proof. Since $C_j(A,\Lambda_+)$ is linear in A, and Λ_+ has finite order, we may take $A = \mathrm{Op}(a,\theta)$ for a symbol a homogeneous of degree r, and patch function θ. Further, as in § 3, we may expand $\mathrm{Op}(a,\theta)$ in a series $\sum_1^\infty A_k S_k$, where A_k consists of multiplication by a bounded function all of whose derivatives are bounded, and S_k is a translation-invariant operator of order $\leqq r$. Since S_k consists of multiplication of Fourier transforms by some function, we have $S_k\Lambda_+ = \Lambda_+ S_k$ and consequently, by induction, $C_j(A_kS_k,\Lambda_+) = C_j(A_k,\Lambda_+)S_k$. Thus $C_j(\sum_1^\infty A_kS_k,\Lambda_+) = \sum_1^\infty C_j(A_k,\Lambda_+)S_k$, and we are reduced to showing that $C_j(A,\Lambda_+)$ has order zero when A is multiplication by a bounded function, all of whose derivatives are bounded. (We are neglecting the convergence of the series, but this follows from the differentiability of the symbol a.)

Suppose A is multiplication by ϕ, and let $A^{(k)}$ denote multiplication by $(\partial/\partial t)^k\phi$. Define an operator Λ_n by $(\Lambda_n f)^\wedge(\xi,\tau) = (1 + |\xi|^2)^{1/2} f^\wedge(\xi,\tau)$; then we have $\Lambda_+ = \Lambda_n - \partial/\partial t$. Let $C_0(A,\Lambda_n) = A$ and

$$C_{j+1}(A,\Lambda_n) = C_j(A,\Lambda_n)\Lambda_n - \Lambda_n C_j(A,\Lambda_n).$$

Then we find that

$$C_j(A^{(k)},\Lambda_n)(-\partial/\partial t) - (\partial/\partial t)C_j(A^{(k)},\Lambda_n) = C_j(A^{(k+1)},\Lambda_n),$$

and hence

$$C_k(A,\Lambda_+) = \sum_0^k C_j^k C_j(A^{(k-j)},\Lambda_n),$$

with C_j^k the combinatorial coefficients. Thus it is sufficient to obtain bounds on the terms of this last sum.

Now we may view $C_j(A^{(k-j)},\Lambda_n)$ as a one-parameter family of operators $C_j(A^{(k-j)}{}_t,\Lambda_n)$, where $A^{(k-j)}$ denotes the operator on $\mathcal{S}(R^n)$ given by $(A^{(k-j)}{}_t g)(x) = (\partial/\partial t)^{k-j}\phi(x,t)g(x)$, and Λ_n applies now to $\mathcal{S}(R^n)$. Thus

$C_j(A^{(k-j)},\Lambda_n)f(x,t)=C_j(A^{(k-j)}{}_t,\Lambda_n)f_t(x)$ with $f_t(x)=f(x,t)$. It is clear then that

$$\| C_j(A^{(k-j)},\Lambda_n)\|_{n+1}\leqq \max_t \| C_j(A^{(k-j)}{}_t,\Lambda_n)\|_n, \tag{20}$$

where $\|\cdot\|_n$ denotes the norm of an operator on $L^p(R^n)$. Further, since Λ_n is a CZO of order -1 on $\mathcal{S}(R^n)$, $\| C_j(A^{(k-j)}{}_t,\Lambda_n)\|_n$ is finite and can be dominated by the supremum of the derivatives of ϕ of order $\leqq k+1$ (see the remark at the end of this proof); thus $\| C_j(A^{(k-j)},\Lambda_n)\|_{n+1}$ can be so dominated. We obtain bounds on $L_s{}^p$ for $s=1,2,\cdots$ by noting that $\Lambda_+{}^s$ is an isomorphism of $L_s{}^p$ onto $L_0{}^p$, and that for s an integer $\geqq 1$.

$$\Lambda_+{}^{-s}C_j(A,{}^{(k-j)}\Lambda_+)\Lambda_+{}^s=\sum_{l=0}^{s} C_l{}^s C_{j+l}(A,{}^{(k-j)}\Lambda_+).$$

Thus the norm on $L_s{}^p$ is dominated by derivatives of ϕ of order $\leqq k+s+1$ for $s=0,1,\cdots$. Like bounds for $s=-1,-2,\cdots$ are obtained by taking adjoints, and then for all real s by interpolating.

This concludes the proof of Lemma 10.

Remark. None of the articles on CZO's (or pseudo-differential operators) gives an explicit account of what derivatives suffice to dominate $\| C_j(B,\Lambda_n)\|_n$, as required in the proof above. The bound we have given can be obtained with the methods of reference [7], while bounds involving various other finite numbers of derivatives follow from a careful accounting of the arguments of the other references. In the case of [8] and [9] one must also assume that the function ϕ, by which A multiplies, vanishes at ∞ in an appropriate way. In any case, the accounting is not very interesting.

Now we return to the proof of Lemma 7, carrying on the outline of the three paragraphs preceding Lemma 9. By induction on k we show that if $u\in L_s{}^p(R^{n+1})$ for $s>-1-k+1/p$, and u vanishes in $\{t>0\}$, then

$$\lim_{t\to 0+}(D_t{}^lBu)(\cdot,t) \text{ exists in } B_{s+m-l-1/p}{}^p(R^n), \tag{21}$$

and has a norm dominated by the norm of u in $L_s{}^p$, for each integer $l\geqq 0$. The induction is based on the formula

$$Bu=C_k(B,\Lambda_+)\Lambda_+{}^{-k}u+\sum_{j=1}^{k}\gamma_{jk}\Lambda_+{}^jB\Lambda_+{}^{-j}u \tag{22}$$

with appropriate constants γ_{jk}. Since $\Lambda_+{}^{-k}$ has order $-k$, Lemma 10 and the restriction theorem show that the first term on the right in (22) has derivatives of order $<m$ with boundary values in $B_{s+m+k-l-1/p}{}^p$. By induction, using Lemma 9,

$$\lim_{t\to 0+}(D_t B\Lambda_+^{-j}u)(\cdot,t)$$

exists in $B_{s+j+m-l-1/p}{}^p$ for all l; thus by the special form of $\Lambda_+ = \Lambda_n - \partial/\partial t$

$$\lim_{t\to 0+}(D_t \Lambda_+^{j} B\Lambda_+^{-j}u)(\cdot,t)$$

exists in $B_{s+m-l-1/p}{}^p$ for all l. Thus from (22), (21) holds with u in $L_s{}^p(R^{n+1})$ and $l < m$, and hence (21) holds for all m. This proves Lemma 7, and with it Theorem 6.

8. Generalizations. Let G_+ be a compact manifold with boundary X, and E and F be vector bundles over G, the double of G_+. Then, as in [11] or [13], we may consider CZO's mapping sections of E into sections of F. Again the top-order symbol is defined; if η_x is a cotangent vector at $x \in G$, then $\sigma(A)(\eta_x)$ is a homomorphism of E_x into F_x, the fibres over x. A is elliptic at x if $\sigma(A)(\eta_x)$ is invertible for each η_x. The spaces $L_s{}^p(E)$, $B_{s-1/p}{}^p(E \mid X)$, and $L_s{}^p(E \mid G_+)$ can also be defined. Here $E \mid G_+$ and $E \mid X$ denote the restrictions of E, and $L_s{}^p(E \mid G_+)$ consists of those sections of $E \mid G_+$ having an extension in $L_s{}^p(E)$. Assume that A is a CZO of order m, that A is elliptic in a neighborhood of G_+, and that near X A is a differential operator. Let $K(A,s,p) = \{u \in L_s{}^p(E \mid G_+) : Au = 0$ in the interior of $G_+\}$. Let $N_0(A) = \{u \in L_\infty{}^p(E) : Au = 0$ and $u = 0$ off $G_+\}$. (Note that $L_\infty{}^p(E)$ is the C^∞ sections of E.) Finally, introduce near X a fixed representation of E as $(E \mid X) \times \{-1 < t < 1\}$, so that $R_\epsilon f$ is defined for $f \in L_s{}^p(E)$ when $s > m - 1 + 1/p$. Let $B_{(s)}{}^p$ denote $\bigoplus_0^{m-1} B_{s-1/p-j}{}^p(E \mid X)$.

THEOREM 7. *$N_0(A)$ is finite dimensional. For $u \in K(A,s,p)$, the $\lim R_\epsilon u = R_{0+}u$ exists in $\bigoplus_0^{m-1} B_{s-1/p-j}{}^p(E \mid X)$. There is a continuous map P of $\bigoplus_0^{m-1} B_{s-1/p-j}{}^p(E \mid X)$ into $K(A,s,p)$, such that*

$$K(A,s,p) = P(\oplus B_{s-1/p-j}{}^p) \oplus N_0(A).$$

$P^+ = R_{0+}P$ is a CZO, and a projection of $B_{(s)}{}^p$ onto $R_{0+}(K(A,s,p))$. The symbol of P^+ can be described by reducing A to a first order system, as in the end of § 6.

This result can be reduced to Theorems 5 and 6 with a technique suggested by Calderón. The matrix $\begin{pmatrix} A & 0 \\ 0 & A^* \end{pmatrix}$ of operators is a self-adjoint operator on $E \oplus F$, and is elliptic in $\{t \geqq -\epsilon\}$ for some $1 > \epsilon > 0$. Because a self-adjoint matrix is homotopic to the identity in a canonical way, there is a C^∞ elliptic symbol F such that $F=\sigma\begin{pmatrix} A & 0 \\ 0 & A^* \end{pmatrix}$ in $\{t \geqq -\epsilon\}$ and $F(\eta_x)$ is a multiple of the identity for x in $\{t \leqq -1\}$. Let B be a CZO with $\sigma(B)=F$, and ϕ and ψ be C^∞ functions with $\phi^2+\psi^2=1$ in $\{t \geqq -\epsilon/2\}$, $\psi=1$ in $\{t \leqq -\epsilon\}$. Then $B^\sim = M_\phi \begin{pmatrix} A & 0 \\ 0 & A^* \end{pmatrix} M_\phi + M_\psi B M_\psi$ is elliptic, and agrees with $\begin{pmatrix} A & 0 \\ 0 & A^* \end{pmatrix}$ on $G_+ \cup X$. From this follows the claim that $N_0(A)$ is finite dimensional, for the map $f \to f \oplus 0$ is an injection of $N_0(A)$ into the null space of $B^\sim$, which is finite dimensional.

Now let N denote the orthogonal projection in $L^2(E \oplus F)$ onto $\{f \oplus g : B^\sim(f \oplus g)=0$ and $f \oplus g = 0$ in $G_-\} = N_0(A) \oplus N_0(A^*)$. (We assume an appropriate Hermitian inner product in E and F, and volume element on G.) As a subspace of the null space of $B^\sim$, the range of N is spanned by finitely many C^∞ sections, so N has order $-\infty$. Next let θ be a real C^∞ function vanishing precisely on $G_+ \cup X$. Then $A^\sim = B^\sim + N + iM_\theta$ is invertible, and Theorems 5 and 6 apply to $A^\sim$; the proofs were carried out for functions, but they apply to operators on bundles with changes in notation only. Let $P^\sim$ and $P^{\sim+}$ be the maps associated with $A^\sim$ whose existence is proved in Theorem 5. These may be represented as 2×2 matrices, and we let P and P^+ be the first entry in the corresponding matrix; thus $Pf=$ the first component of $P^\sim(f \oplus 0)$.

It is clear that P maps $\overset{m-1}{\underset{0}{\oplus}} B_{s-j-1/p}{}^p$ continuously into $L_s{}^p(E \mid G_+)$, and since $A^\sim P^\sim{}_s(f \oplus 0)=0$ inside G_+, it follows that Pf is orthogonal to $N_0(A)$ and $APf=0$ inside G_+. In particular, P maps into $K(A,s,p)$.

To obtain the direct sum for $K(A,s,p)$, let $u \in K(A,s,p)$, and u be orthogonal to $N_0(A)$. Then $u \oplus 0 \in K(A^\sim,s,p)$, so $\lim\limits_{\epsilon \to 0+} R_\epsilon(u \oplus 0)$ exists and lies in the range of $P^{\sim+}$, and hence by Lemma 7 $u \oplus 0 = P^\sim \lim R_\epsilon(u \oplus 0)$. Thus $u = P \lim R_\epsilon u$ when u is orthogonal to $N_0(A)$, which with the previous paragraph shows $K(A,s,p)$ is the (orthogonal) direct sum

$$P(\overset{m-1}{\underset{0}{\oplus}} B_{s-j-1/p}{}^p(E \mid X)) \oplus N_0(A).$$

Since $R_{0+}P^\sim$ is a CZO, the same holds for $P^+ = R_{0+}P$, and the symbol of the latter can be derived from that of $R_{0+}P^\sim$. Because the range of P lies in $K(A,s,p)$, the range of P^+ lies in $R_{0+}(K(A,s,p))$. Conversely, if $f=R_{0+}u$ for a u in $K(A,s,p)$, then $v=u$ minus its orthogonal projection on $N_0(A)$; for $R_{0+}\phi=0$ when $\phi \in N_0(A)$. Since $v=PR_{0+}v=Pf$ as in the previous paragraph, we have $P^{+}f = R_{0+}Pf = R_{0+}v = f$. This identifies the range of P^{+} as in Theorem 7, and shows that P^{+} is a projection.

Finally, the fact that $R_{0+}(K(A,s,p))$ is the closure of $R_{0+}(K(A,\infty,p))$ is proved as in Theorem 5.

BRANDEIS UNIVERSITY.

REFERENCES.

[1] S. Agmon, A. Douglis and L. Nirenberg, "Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions, II," *Commentarii of Pure and Applied Mathematics*, vol. 17 (1964), pp. 35-92.

[2] F. Browder, "Estimates and existence theorems for elliptic boundary value problems," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 45 (1959), pp. 365-372.

[3] ———, "A priori estimates for solutions of elliptic boundary value problems," *Nederlandse Akademie van Wetenschappen*, vol. 22 (1960), pp. 145-169, and vol. 23 (1961), 404-410.

[4] A. P. Calderón, "Lebesgue spaces of differentiable functions and distributions," *Proceedings of Symposia in Pure Mathematics*, vol. 4, A.M. S., Providence.

[5] ———, "Intermediate spaces and intepolation," Warsaw Conference, September, 1960.

[6] ———, "Boundary value problems for elliptic equations," *Outlines of the Joint Soviet-American Symposium on Partial Differential Equations*, August, 1963. Novosibirsh.

[7] A. P. Calderón and A. Zygmund, "Singular integral oueraters and differential equations," *American Journal of Mathematics*, vol. 79 (1957), pp. 901-921.

[8] L. Hörmander, "Pseudo-differential operators," *Commentarii of Pure and Applied Mathematics*, to appear.

[9] J. J. Kohn and L. Nirenberg, "An algebra of pseudo-differential operators," *Commentarii of Pure and Applied Mathematics*, vol. 18 (1965), pp. 269-305.

[10] J. L. Lions, "Une construction d'espaces d'intepolation," *Comptes Rendus de l'Academie des Sciences*, Paris, vol. 251 (1960), pp. 1851-1855.

[11] R. Palais, *Seminar on the Atiyah-Singer index theorem*, Annals of Mathematics Studies, 1965.

[12] R. Seeley, "Refinement of the functional calculus of Calderón and Zygmund," *Proceedings, Koniklijke Nederlandse Akademie van Wetenschappen*, vol. 68 (1965), pp. 521-531.

[13] ———, "Singular integral operators on vector bundles," *Transactions of the American Mathematical Society*.

[14] E. M. Stein, "The characterization of functions arising as potentials, II," *Bulletin of the American Mathematical Society*, vol. 68 (1962), pp. 577-582.

[15] M. H. Taibleson, "Lipschitz classes of functions and distributions in E_n," *Bulletin of the American Mathematical Society*, vol. 69 (1963), pp. 487-493.

[16] A. Zygmund, *Trigonometrical series*, 1935.

THE SPECTRA OF GENERALIZED TOEPLITZ OPERATORS.

By Lewis A. Coburn.

1. Introduction. Recently, Devinatz [2] extended a major portion of the theory of Toeplitz operators acting on the usual H^2-space of the circle to a similarly defined class of operators on some more general H^2-spaces. The H^2-spaces in question are associated with certain algebras of complex-valued continuous functions on compact Hausdorff spaces. In fact, it is not hard to see that Devinatz's general results hold in the logmodular algebra setting described by Hoffman [4] where X is a compact Hausdorff space, A is a logmodular algebra of complex-valued functions on X, μ is a representing measure for A and $H^2 \equiv H^2(\mu, A, X)$ is defined to be the closure of A in $L^2(\mu, X)$. Then for ϕ in $L^\infty(\mu, X)$, P the orthogonal projection from L^2 onto H^2, and f in H^2 the Toeplitz operator T_ϕ is defined by

$$T_\phi f = P(\phi \cdot f).$$

For the circle case (X the unit circle, μ normalized Lebesgue measure on X, A the uniformly closed algebra generated by 1 and z), precise information has been obtained about the spectrum of T_ϕ when ϕ is continuous or equivalent a. e. to a continuous function. Although the "winding number theorem" of Calderon, Spitzer, and Widom [1] in the circle case cannot be fully generalized, a good deal can be said about the spectrum $\sigma(T_\phi)$ in the logmodular algebra setting when ϕ is continuous on X (or when ϕ is equal a. e. (μ) to a continuous function). In this note, I give a generalization of a weakened version of the "winding number theorem" which holds when A is a Dirichlet algebra [3] but fails for certain non-Dirichlet logmodular algebras. It is interesting that as a consequence of this "generalized winding number theorem," results about the algebra A can be obtained. A particular consequence of the main theorem is the result that for A a Dirichlet algebra on X and μ any representing measure for A, the support of μ [5] is connected.

In Section 2 of this paper, some basic properties of generalized Toeplitz operators are stated and a preliminary result established. Section 3 contains the proof of the main theorem and some consequences. These results appeared

---

Received October 14, 1964.
Revised May 6, 1965.

in my doctoral dissertation at the University of Michigan and in this connection it is a pleasure to thank Professors P. R. Halmos and R. G. Douglas for their helpful advice.

2. Preliminary results. We begin by stating some facts about Toeplitz operators in the logmodular algebra setting. Proofs either appear in [2] or follow the standard proofs in the circle case. If K is a subset of the complex plane, $\mathscr{C}(K)$ denotes the convex hull of K. The essential supremum norm of a function ϕ in L^∞ is denoted by $\|\phi\|_\infty$. As usual, the sup. norm algebra H^∞ is defined by $H^\infty = L^\infty \cap H^2$ and $\overline{H^\infty}$ denotes the algebra of complex conjugates of functions in H^∞. For ϕ in L^∞, the Laurent operator associated with ϕ is defined on f in L^2 by

$$L_\phi f = \phi \cdot f.$$

It is well-known that $\sigma(L_\phi)$ is just the essential range of ϕ.

LEMMA 1. *The inclusions of* $\sigma(L_\phi) \subset \sigma(T_\phi) \subset \mathscr{C}[\sigma(L_\phi)]$ *hold for all* ϕ *in* L^∞.

LEMMA 2. *For* ϕ *in* L^∞, $\|T_\phi\| = \|L_\phi\| = \|\phi\|_\infty$.

LEMMA 3. *The equality* $T_\phi T_\psi = T_{\phi\psi}$ *holds if either* $\psi \in H^\infty$ *and* $\phi \in L^\infty$ *or* $\bar{\phi} \in H^\infty$ *and* $\psi \in L^\infty$.

LEMMA 4. *Suppose* ϕ *and* $1/\phi$ *are in* L^∞ *and* $\theta = \phi/|\phi|$. *Then* T_ϕ *is invertible if and only if* T_θ *is invertible.*

LEMMA 5. *For* ϕ *real-valued in* L^∞, $\sigma(T_\phi)$ *is the closed interval* $[\text{ess. inf. } \phi, \text{ess. sup. } \phi]$.

In what follows, $\operatorname{Re} A$ denotes the collection of real parts of functions in the logmodular algebra A and we write $\exp \phi$ for the composition function $e^z \circ \phi$. The unit circle (perimeter) in the complex plane is denoted by T and the closed unit disc is denoted by D.

THEOREM 1. *If* θ *is in the uniform closure of* $\operatorname{Re} A$ *then* $\sigma(T_{\exp i\theta}) = \sigma(L_{\exp i\theta})$.

Proof. Clearly

$$\sigma(L_{\exp i\theta}) = \text{ess. range}(\exp i\theta) \subset T.$$

By Lemma 1, $\sigma(T_{\exp i\theta}) \subset \mathscr{C}[\sigma(L_{\exp i\theta})]$ and so $\sigma(T_{\exp i\theta}) \subset D$. Suppose that $|\lambda| < 1$ and consider $T_{\exp i\theta} - \lambda$. Since θ is in the uniform closure of $\operatorname{Re} A$, there is a ϕ in A with $\theta - \phi - \bar{\phi} = \psi$ and

$$\| \psi \|_\infty < \ln[2 - |\lambda|].$$

Now using Lemma 3 and he fact that the Banach algebras H^∞ and $\bar{H}^\infty$ are closed under exponentiation, we see that

$$T_{\exp i\theta} - \lambda = T_{\exp i\bar{\phi}} T_{\exp i\psi} T_{\exp i\phi} - \lambda.$$

Again using Lemma 3 and the standard power-series definition of the exponential, we see that $T_{\exp i\phi}$ and $T_{\exp i\bar{\phi}}$ are invertible with inverses $T_{\exp -i\phi}$ and $T_{\exp -i\bar{\phi}}$ respectively. It follows that

$$(*) \qquad T_{\exp -i\bar{\phi}} T_{(\exp i\theta) - \lambda} T_{\exp -i\phi} = T_{(\exp i\psi) - \lambda \exp -i(\phi + \bar{\phi})}.$$

Using Lemma 2 and expanding in power series, we now find that

$$1 > (\exp \| \psi \|_\infty) - 1 + |\lambda| \geqq \| (\exp i\psi) - 1 \|_\infty + |\lambda|$$
$$\geqq \| (\exp i\psi) - \lambda \exp - i(\phi + \bar{\phi}) - 1 \|_\infty \geqq \| T_{(\exp i\psi) - \lambda \exp -i(\phi + \bar{\phi})} - 1 \|.$$

It follows from a standard result in operator theory that $T_{(\exp i\psi) - \lambda \exp -i(\phi + \phi)}$ is invertible. Thus, by (*), $T_{\exp i\theta} - \lambda$ is also invertible and so we have the result that

$$\sigma(T_{\exp i\theta}) \subset \mathscr{C}[\sigma(L_{\exp i\theta})] \cap T.$$

It follows that $\sigma(T_{\exp i\theta}) \subset \sigma(L_{\exp i\theta})$ and since the reverse inclusion was established generally in Lemma 1, the proof is completed.

3. Main result. Before stating the main result, we need to examine a geometrical notion similar to convexity. A subset K of the complex plane is *ray-convex* if for any point p not in K there is a ray from p which does not intersect K. The above definition can be rephrased in the existing language of complex variables as: the complement of K is star-shaped with respect to the point at infinity. In any case, it is clear that the intersection of any family of ray-convex sets is again ray-convex. Thus, for any subset of the plane K, there is a smallest ray-convex set containing K. We denote this set by $\mathscr{R}(K)$. By comparing the definitions of ray-convexity and convexity, it is easy to see that for any planar subset K the relation $K \subset \mathscr{R}(K) \subset \mathscr{C}(K)$ holds.

Now let S_μ be the (closed) support set of the Baire meausure μ on X.

THEOREM 2. *Let A be a Dirichlet algebra on X and suppose that μ is a representing measure for A. If ϕ is continuous on S_μ then*

$$\phi(S_\mu) \subset \sigma(T_\phi) \subset \mathscr{R}[\phi(S_\mu)].$$

Proof. Without loss of generality we may assume that $S_\mu = X$. Otherwise, cutting the elements of A down to S_μ and taking the uniform closure in $C(S_\mu)$ we get a new Dirichlet algebra A' with μ as a representing measure supported on all of S_μ. It is easy to see that $\sigma(T_\phi)$ is not changed by this process since $H^2(\mu, A, X) = H^2(\mu, A', S_\mu)$.

Now taking $S_\mu = X$, it is clear that ϕ is in $L^\infty(\mu)$ with $\phi(X) =$ ess. range ϕ. Hence, by Lemma 1, $\phi(X) \subset \sigma(T_\phi)$. Suppose that λ is not a point of $\mathfrak{R}[\phi(X)]$. Then from the definition of ray-convevity it follows that the function

$$\frac{\phi(x) - \lambda}{|\phi(x) - \lambda|}$$

is continuous on X and the image, Y, of X under this function omits a complex number of modulus one, $e^{i\theta}$ with $0 \leqq \theta < 2\pi$. In fact, since Y is compact, Y must omit an open disc about $e^{i\theta}$. Let $g(x)$ be the unique real number in the half-open interval $[\theta, \theta + 2\pi)$ such that

$$z = e^{ig(z)}$$

for z in Y. It is easy to check that $g(z)$ is continuous on Y. Since the composition of continuous functions is continuous, it follows that

$$\rho(x) = g\left[\frac{\phi(x) - \lambda}{|\phi(x) - \lambda|}\right]$$

is continuous and real-valued on X with

$$\frac{\phi(x) - \lambda}{|\phi(x) - \lambda|} = e^{i\rho(x)}.$$

Hence, using Lemma 4, we see that $T_\phi - \lambda$ is invertible if $T_{\exp i\rho}$ is invertible. Since A is a Dirichlet algebra, ρ is automatically in the uniform closure of $\operatorname{Re} A$ and so an application of Theorem 1 completes the proof.

COROLLARY 1. *If A is a Dirichlet algebra on X and μ is a representing measure for A then S_μ is connected.*

Proof. If S_μ is not connected then there are non-empty disjoint closed subsets E and F in X such that $S_\mu = E \cup F$. Since X is normal, there is a continuous real-valued function ϕ on X with $\phi(E) = \{1\}$ and $\phi(F) = \{0\}$. Hence

$$\phi(S_\mu) = \mathfrak{R}[\phi(S_\mu)] = \{0, 1\}$$

so by Theorem 2 we must have $\sigma(T_\phi) = \{0, 1\}$. This contradicts Lemma 5.

It is natural to ask whether Theorem 2 extends to the case where A is a logmodular algebra, but not a Dirichlet algebra on X. To this end, consider the algebra L^∞ in the circle and denote its maximal ideal space by $\mathfrak{M}(L^\infty)$. It is well-known that under the Gelfand transform, H^∞ is isometrically isomorphic to a logmodular algebra A' on $\mathfrak{M}(L^\infty)$. Further, $\mathfrak{M}(L^\infty)$ is known to be totally disconnected [3]. Thus, if Theorem 2 and consequently Corollary 1 hold, we see that each representing measure for A' must be concentrated at a point of $\mathfrak{M}(L^\infty)$. From this, we obtain a homeomorphism

$$\mathfrak{M}(L^\infty)) \cong \mathfrak{M}(H^\infty)$$

which is absurd since it well-known that $\mathfrak{M}(H^\infty)$ is connected. Other counterexamples can be obtained by considering the logmodular algebras described in [4].

Finally, we note that Corollary 1 furnishes a new proof of a well-known result in the theory of topological groups [5].

THEOREM 3. *If G is a compact abelian group with linearly ordered dual $\hat{G}$ then G must be connected.*

Proof. Let A be the uniformly closed algebra of continuous functions on G generated by the non-negative elements in the order on $\hat{G}$. It is easy to see that A is a Dirichlet algebra on G and that Haar measure μ on G is a representing measure for A. Thus, by Corollary 1, $S_\mu = G$ is connected.

REFERENCES.

[1] A. Calderon, F. Spitzer and H. Widom, "Inversion of Toeplitz matrices," *Illinois Journal of Mathematics*, vol. 3 (1959), pp. 490-498.

[2] A. Devinatz, "Toeplitz operators on H^2 spaces," *Transactions of the American Mathematical Society*, vol. 112 (1964), pp. 304-317.

[3] K. Hoffman, *Banach spaces of analytic functions*, Englewood Cliffs, N. J., 1962.

[4] ———, "Analytic functions and logmodular Banach algebras," *Acta Mathematica*, vol. 108 (1962), pp. 271-317.

[5] W. Rudin, *Fourier analysis on groups*, New York, 1962.

THE KERVAIRE INVARIANT OF $8k+2$-MANIFOLDS.

By Edgar H. Brown, Jr. and Franklin P. Peterson.[1]

1. Introduction. The main results of this paper were announced in [6]. Let $\Omega_n(e)$, $\Omega_n(SU)$, and $\Omega_n(Spin)$ denote the n-th framed, SU, and $Spin$ cobordism groups respectively (see [7] and [11].). In [8] Kervaire defined a homomorphism $\Phi\colon \Omega_{4k+2}(e) \to Z_2$, $k \neq 0, 1, 3$, which is the obstruction to a framed $4k+2$-manifold being framed cobordant to a homotopy sphere ([9]). Kervaire showed that $\Phi = 0$ for $k = 2, 4$. In [4] a homomorphism ψ: $\Omega_{8k+2}(Spin) \to Z_2$ was defined such that $\Phi = \psi\rho$ where $\rho\colon \Omega_n(e) \to \Omega_n(Spin)$ is the obvious map. The obvious map of $\Omega_n(SU)$ into $\Omega_n(Spin)$ defines a homomorphism of $\Omega_{8k+2}(SU)$ into Z_2 which we also denote by ψ. This latter map is the main object to the investigated in this paper. In an appendix we briefly discuss $\psi\colon \Omega_{8k+2}(Spin) \to Z_2$.

The main results of this paper are as follows:

Theorem 1.1. *$\Phi\colon \Omega_{8k+2}(e) \to Z_2$ is zero for $k > 0$.*

The following corollaries of (1.1) are implied by the results of [9], [8] and [3].

Corollary 1.2. *$bP_{8k+2} \approx Z_2$, where bP_{8k+2} is the group of homotopy spheres which bound stably parallelizable $8k+2$-manifolds* [9].

Corollary 1.3. *If K is the topological manifold obtained by plumbing two copies of the tangent disc bundle of S^{4k+1} together and then attaching an $8k+2$-disc, then K does not admit a diffedentiable structure.*

(1.3) follows from [8] if one has the result that a C^∞ manifold with underlying topological space K is stably parallelizable. In Appendix 2 we give a proof of this due to John Milnor.

Corollary 1.4. *Every element of $\Omega_{8k+2}(e)$ can be represented by a homotopy sphere, $k \geqq 1$.*

Corollary 1.5. *A finite, 1-connected CW complex has the homotopy type of a stably parallelizable $8k+2$-manifold if and only if there is a stably*

Received June 7, 1965.

[1] The first named author was partially supported by the N. S. F. and the second named author was partially supported by the N. S. F. and by the U. S. Army Research Office.

spherical class $m \in H_{8k+2}(X)$ *such that* $m \cap : H^q(X;Z) \approx H_{8k+2-q}(X;Z)$ *for all* q *and* $\psi(X) = 0$ *(see* § 3 *for a definition of* $\psi(X)$*)* ([3], [12]).

It is known that $\Omega_2(SU) \approx Z_2$ [7]. Let α be the generator. Define $\psi\colon \Omega_2(SU) \to Z_2$ by $\psi(\alpha) = 1$.

THEOREM 1.6. *If* $\beta \in \Omega_{8k+2}(SU)$ *and* $\gamma \in \Omega_{8l}(SU)$, $k \geqq 0$, $l > 0$, *then*

$$\psi(\beta\gamma) = \psi(\beta) I(\gamma)$$

where $I(\gamma)$ *is the index of* γ *reduced* mod 2. *(*$I(\gamma)$ *= Euler characteristic of* γ mod 2 *also.)*

In [7] it is shown that $\Omega_{16}(SU)$ contain an element γ^{16} with $I(\gamma^{16}) \neq 0$ mod 2.

By (1.6) we have:

COROLLARY. 1.7. $\psi(\alpha\gamma^{16}) = 1$.

In § 2 we give some preliminary results about cohomology operations. In § 3 we define $\psi\colon \Omega_{8k+2}(SU) \to Z_2$ and in § 4 we prove Theorems (1.1) and (1.6).

2. Some cohomology operations.

Throughout the remainder of the paper all homology and cohomology groups will have Z_2 coefficients unless otherwise specified. m will denote an integer of the form $4k+1$, $k > 0$. Below we show that various cohomology operations are equal. This will always mean equal modulo the largest indeterminacy involved.

In [5] it is shown that the relation

$$Sq^2 Sq^{m-1} + Sq^1(Sq^2 Sq^{m-2}) = 0$$

on m-dimensional cohomology classes gives rise to a secondary operation

$$\begin{aligned}\phi\colon & H^m(X) \cap \operatorname{Ker} Sq^{m-1} \cap \operatorname{Ker} Sq^2 Sq^{m-2} \\ & \to H^{2m}(X)/Sq^1 H^{2m-1}(X) + Sq^2 H^{2m-2}(X).\end{aligned} \tag{2.1}$$

Furthermore, the following is proved:

2.2) If $\phi(u)$ and $\phi(v)$ are defined, $\phi(u+v)$ is defined and

$$\phi(u+v) = \phi(u) + \phi(v) + uv$$

Let $\hat{\sigma}_m \in H^m(K(Z,m);Z)$ be the generator and let σ_m be $\hat{\sigma}_m$ reduced mod 2. Let $p\colon E \to K(Z,m)$ be the fibration with fibre $K(Z_2, 2m-2)$ and k-invariant $Sq^{m-1}\hat{\sigma}_m$. Then $\phi(p^*\sigma_m)$ is defined since, by the Adem relations,

$$Sq^2 Sq^{m-2} = Sq^m + Sq^{m-1} Sq^1 = Sq^1 Sq^{m-1} + Sq^{m-1} Sq^1$$

is zero on $p^*\sigma_m$. Choose an element $z \in \phi(p^*\sigma_m)$. z defines a cohomology operation

$$\hat{\phi}\colon H^m(X,Z) \cap \operatorname{Ker} Sq^{m-1} \rightarrow H^{2m}(X)/Sq^2H^{2m-2}(X)$$

The following is immediate.

(2.3) If $\hat{u} \in H^m(X;Z)$, $Sq^{m-1}\hat{u} = 0$ and u is $\hat{u}$ reduced mod 2, then

$$\hat{\phi}(\hat{u}) = \phi(u).$$

Let $\hat{u} \in H^m(X;Z)$ be viewed as a map $\hat{u}\colon X \rightarrow K(Z,m)$. Then the following is proved in [13].

(2.4) If $Sq^{m-1}\hat{u} = 0$,

$$\hat{\phi}(\hat{u}) = Sq^2{}_{\hat{u}}(Sq^{m-1}\hat{\sigma}_m)$$

Let $f\colon (X,A) \rightarrow (Y,B)$ and let $g\colon X \rightarrow Y$ be the map defined by f. We need a product formula for functional operations.

(2.5) If $u \in H^q(Y,B)$, $v \in H^p(Y)$ and $g^*v = Sq^2v = Sq^1u = Sq^2u = 0$, then

$$Sq^2{}_f(uv) = (f^*u)(Sq^2{}_gv)$$

(This formula is proved, in the absolute case, in [1].)

Proof. Let $h\colon A \rightarrow B$ be the map defined by f. Note that (X,A) is contained in the mapping cylinders (C_g, C_h) and that $A = C_h \cap X$. Hence we may assume f is an inclusion map and $A = X \cap B$. Cup product with u maps the exact sequence of (Y,X) into the exact sequence of the triad (Y,X,B) giving the following commutative diagram:

$$\begin{tikzcd}
H^{p-1}(X) \arrow[r] \arrow[d, "\cup f^*u"] & H^p(Y,X) \arrow[r] \arrow[d, "\cup u"] & H^p(Y) \arrow[r] \arrow[d, "\cup u"] & H^p(X) \arrow[d] \\
H^{p+q-1}(X,A) \arrow[r] & H^{p+q}(Y,X \cup B) \arrow[r] & H^{p+q}(Y,B) \arrow[r] & H^{p+q}(X,A)
\end{tikzcd}$$

Since $Sq^1u = Sq^2u = 0$, $Sq^2uz = uSq^2z$ for any z. Hence Sq^2 maps the above ladder into itself giving a large commutative diagram. (2.5) now follows by chasing around this diagram.

3. Definition of the Kervaire invariant. Throughout this section we view classes $u \in H^q(X)$ as maps $u\colon X \rightarrow K(Z_2,q)$. Again, $m = 4k+1$, $k > 0$.

$\Omega_n(X;SU)$ will denote the n-th SU bordism group [7]. Recall, this group is the set of equivalence classes, under an appropriately defined cobordism relation, of triples (M, λ, f) where M is a closed, compact C^∞ n-manifold, λ is an SU reduction of the normal bundle of M embedded in R^{n+k} for large k and $f\colon M \rightarrow X$. $\Omega_n(SU) = \Omega_n(pt; SU)$. One may easily show that if X is connected, every element of $\Omega_n(X; SU)$ can be represented by (M, λ, f) where M is connected. Hereafter we assume all spaces are connected. Also we assume all manifolds have an SU structure on their normal bundle, (M, λ, u) will be denoted by (M, u) and $\nu_M\colon M \rightarrow BSU_k$ will denote the map defined by this SU structure.

LEMMA 3.1. *If* $\{M, u\} \in \Omega_{2m}(K(Z_2, m); SU)$, *then* $\{M, u\} = \{M', u'\}$ *where* M' *is* 1-*connected. Furthermore, there is a cobordism* (N, v) *between* (M, u) *and* (M', u') *such that if* $i\colon M \rightarrow N$ *and* $j\colon M' \rightarrow N$ *are the inclusion maps,*

$$i^*\colon \mathrm{H}^q(M) \approx H^q(N) \quad \textit{for } q > 2$$
$$j^*\colon H^q(M') \approx H^q(N) \quad \textit{for } q \neq 2m-1, 2m-2.$$

Proof. We form M' by killing $\pi_1(M)$ by surgery [10]. This process yields a manifold N with an SU reduction such that $\partial N = M - M'$. Furthermore N consists of $M \times I$ with handles $D^2 \times D^{2m-1}$ attached by maps $h\colon S^1 \times D^{2m-1} \rightarrow M \times \{0\}$. Up to homotopy type, N is M with 2-cells attached and N is also M' with $(2m-1)$-cells attached. (3.1) now follows from these properties of N.

Let

$$\bar{\phi}\colon \Omega_{2m}(K(Z_2, m); SU) \rightarrow Z_2$$

be defined as follows: Let ϕ be the cohomology operation described in (2.1). Let $\{M, u\} \in \Omega_{2m}(K(Z_2, m); SU)$. Since M has an SU reduction, its Stiefel-Whitney classes w_1 and w_2 are zero. Therefore by the Wu formulas,

$$Sq^1 H^{2m-1}(M) = w_1 H^{2m-1}(M) = 0 \text{ and } Sq^2 H^{2m-2}(M) = w_2 H^{2m-2}(M) = 0.$$

Hence, $\phi\colon H^m(M) \cap \operatorname{Ker} Sq^{m-1} \rightarrow H^{2m}(M)$. If $Sq^{m-1}u = 0$, we let

$$\bar{\phi}\{M, u\} = \phi(u)([M])$$

where $[M] \in H_{2m}(M)$ denotes the fundamental class. By (3.1) we may always choose M to be 1-connected and hence so that

$$Sq^{m-1}H^m(M) \subset H^{2m-1}(M) \approx H_1(M) = 0.$$

Therefore $\bar{\phi}$ is defined on all of $\Omega_{2m}(K(Z_2, m); SU)$. We show that it is well defined. Let $\beta \in \Omega_{2m}(K(Z_2, m); SU)$. Choose an $(M_1, u_1) \in \beta$ such that

M_1 is 1-connected. Let (M_2, u_2) be any representative of β such that $Sq^{m-1}u_2 = 0$. We show $\phi(u_2)([M_2]) = \phi(u_1)([M_1])$. Let (N, v) be a cobordism between (M_1, u_1) and (M_2, u_2). By surgery we make N 2-connected. Let $j_i\colon M_i \to N$ be the inclusion maps. $H^{2m-1}(N, M_2) \approx H_2(N, M_1) = 0$ since $H_2(N) = H_1(M_1) = 0$. Therefore $j_2{}^*\colon H^{2m-1}(N) \to H^{2m-1}(M_2)$ is an injection. $j_2{}^*Sq^{m-1}v = Sq^{m-1}u_2 = 0$. Hence $Sq^{m-1}v = 0$ is zero. In a similar way one shows that $j_i{}^*\colon H^{2m}(N) \to H^{2m}(M_i)$ is an injection and hence that $Sq^2Sq^{m-2}v = 0$. Therefore $\phi(v)$ is defined and, since $\phi(u_i) = j_i{}^*\phi(v)$, $\phi(u_1) = 0$ if and only if $\phi(u_2) = 0$. Thus $\bar{\phi}$ is well defined.

LEMMA. 3. 2. *If* $\{M, u\} \in \Omega_{2m}(K(Z_2, m); SU)$, $Sq^{m-1}u = 0$ *and* u *is the reduction* mod 2 *of an integer class* $\hat{u}$, *then*

$$\bar{\phi}\{M, u\} = Sq^2_{\hat{u}}\,(Sq^{m-1}\hat{\sigma}_m)([M])$$

Proof. By (2.3) and (2.4),

$$\phi(u) = \hat{\phi}(\hat{u}) = Sq^2_{\hat{u}}\,(Sq^{m-1}\hat{\sigma}_m).$$

Thus the only thing to check is that the indeterminacy for each of these operations is zero. The indeterminacy of ϕ is $Sq^2H^{2m-2}(M) = w_2H^{2m-2}(M) = 0$. The indeterminacy of $Sq^2{}_u$ is

$$Sq^2H^{2m-2}(M) + \hat{u}^*H^{2m}(K(Z, m))$$

Thus we must show that $v = Sq^{i_1}Sq^{i_2}\cdots Sq^{i_l}u = 0$ if $i_1 + i_2 + \cdots + i_l = m$. By the Wu formulas $v = zu$ where z is a polynomial in the Stiefel-Whitney classes of M. Therefore $z = \nu^*{}_M z'$ for $z' \in H^m(BSU)$. But m is odd and hence $z' \in H^m(BSU) = 0$.

We next define the Kervaire Invariant $\psi\colon \Omega_{2m}(SU) \to Z_2$. Let $\{M\} \in \Omega_{2m}(SU)$. Choose a symplectic basis $\{u_i, v_i \mid i = 1, 2, \cdots, \nu\}$ for $H^m(M)$, that is, $u_1, \cdots, u_\nu$, $v_1, \cdots, v_\nu$ is a basis for $H^m(M)$, $u_iu_j = v_iv_j = 0$, and $u_iv_j = \delta_{ij}$. Since M is orientable, $u^2 = 0$, $u \in H^m(M)$, and hence such a basis exists. Define

$$\psi\{M\} = \sum_{i=1}^{\nu} \bar{\phi}\{M, u_i\} \cdot \bar{\phi}\{M, v_i\}. \tag{3.3}$$

We show that ψ is a homomorphism and that it is well defined. By (3.1) we may change M by surgery so that it is 1-connected. Furthermore, by (3.1) $\{u_i, v_i\}$ goes over, under this process, to a symplectic basis. Hence we may assume M is 1-connected. By (2.2)

$$\bar{\phi}\{M,\ \}\colon H^m(M) \to Z_2$$

is a quadratic function whose associated bilinear form, namely, cup product, is non-singular. Therefore the right side of 3.3 is independent of the choice of the symplectic basis [2]. One may easily check that the right side of 3.3 is additive with respect to the connected sum operation on manifolds. Thus ψ is a homomorphism and to show that it is well defined it is sufficient to show that the right side of (3.3) is zero if $M = \partial N$. We may make N 2-connected by surgery. Recall, if $j\colon M \to N$ is the inclusion map, $u \in H^m(N)$, $v \in H^m(N)$, then $u \cdot \delta^* v = \delta^*(j^*(u) \cdot v)$, where $\delta^*\colon H^m(M) \to H^{m+1}(N, M)$. From this fact and Poincare duality one may obtain classes $u_i \in H^m(N)$ and $v_i \in H^m(M)$ such that $\{j^* u_i, v_i\}$ is a symplectic basis for $H^m(M)$. $H^{2m-1}(N) \approx H_2(N, M) = 0$ and $H^{2m}(N) \approx H_1(N, M) = 0$. Therefore $\phi(u_i)$ is defined and equals zero. Therefore $\bar{\phi}(\{M, j^* u_i\}) = \phi(j^* u_i)([M]) = j^* \phi(u_i)([M]) = 0$. Thus ψ is well defined.

Remark 3.4. The secondary operation ϕ is not uniquely determined by the relation between primary operations from which it arises, that is, it is only determined up to the addition of a stable primary operation. (3.2) shows that $\psi\colon \Omega_{8k+2}(SU) \to Z_2$ does not depend on the choice of ϕ since for any $\{M\} \in \Omega_{2m}(SU)$, one may choose M 1-connected and with every element of $H^m(M)$ the reduction of an integer class (See §4.).

Let α be the generator of $\Omega_2(SU) \approx Z_2$. We define $\psi\colon \Omega_2(SU) \to Z_2$ by $\psi(\alpha) = 1$.

Finally, to complete the statement of Corollary 1.5, we define $\psi(X) \in Z_2$ when X is a 1-connected, finite CW complex which has a stably spherical class $m \in H_{8k+2}(X; Z)$ such that

$$\cap\, m\colon H^q(X; Z) \approx H_{8k+2-q}(X; Z)$$

for all q. Since $H_{8k+2}(X; Z)$ is generated by a stably spherical class, Sq^i is zero on $H^{8k+2-i}(X)$. Also

$$Sq^{4k} H^{4k+1}(X) \subset H^{8k+1}(X) \approx H_1(X) = 0$$

Therefore ϕ defines a quadratic function

$$\phi\colon H^{4k+1}(X) \to H^{8k+2}(X)$$

Let

$$\psi(X) = \sum_{i=1}^{r} \phi(u_i)(m) \cdot \phi(v_i)(m)$$

where $\{u_i, v_i \mid i = 1, \cdots, r\}$ is a symplectic basis for $H^{4k+1}(X)$.

4. Proofs of Theorems (1.1) and (1.6). We first prove (1.6). Let

$\beta \in \Omega_{8k+2}(SU)$, $k \geqq 0$ and let $\gamma \in \Omega_{8l}(SU)$, $l > 0$. We wish to show that $\psi(\beta\gamma) = \psi(\beta) I(\gamma)$ where $I(\gamma)$ is the index of $\gamma \bmod 2$.

Let $M \in \beta$ and $N \in \gamma$. Applying surgery to M and N we may choose them so that $\nu_{M*}\colon \pi_i(M) \to \pi_i(BSU)$ is an isomorphism for $i < 4k+1$ and $\nu_{N*}\colon \pi_i(N) \to \pi_i(BSU)$ is an isomorphism for $i < 4l$. Then $H^q(M) = 0$ for q odd and $q \neq 4k+1$ and $H^q(N) = 0$ for q odd. Furthermore the elements of $H^{4k+1}(M)$ and $H^{4l}(N)$ are reduction mod 2 of integer classes because $H^{4l+1}(N;Z) \approx H_{4l-1}(N;Z) = 0$ and $H^{4k+2}(M;Z) \approx H_{4k}(M;Z) \approx H_{4k}(BSU;Z)$ which is free abelian. Note $H^{4(k+l)+1}(M \times N) = H^{4k+1}(M) \otimes H^{4l}(N)$.

LEMMA 4.1. *If* $u \in H^{4k+1}(M)$ *and* $v \in H^{4l}(N)$,

$$\begin{aligned} \phi(u \otimes v) &= \phi(u) \otimes v^2 && \text{if } k > 0 \\ &= 0 && \text{if } k = 0 \text{ and } v^2 = 0. \end{aligned}$$

Proof. Let $\hat{u} \in H^{4k+1}(M;Z)$ and $\hat{v} \in H^{4l}(N;Z)$ be classes which give u and v when reduced mod 2. Below we denote $Sq^2{}_f u$ by $Sq^2(f, u)$.

$$\phi(u \otimes v) = \hat{\phi}(\hat{u} \otimes \hat{v}) \tag{4.2}$$

$$= Sq^2(\hat{u} \otimes \hat{v}, Sq^{4(k+l)} \hat{\sigma}_{4(k+l)+1}) \tag{4.3}$$

$$= Sq^2((\hat{u} \times id)(\hat{\sigma}_{4k+1} \otimes \hat{v}), Sq^{4(k+l)} \hat{\sigma}_{4(k+l)+1})$$

$$= Sq^2(\hat{u} \times id, Sq^{4(k+l)}(\hat{\sigma}_{4k+1} \otimes v)) \tag{4.4}$$

$$= Sq^2(u \times id, Sq^{4k} \hat{\sigma}_{4k+1} \otimes v^2) \tag{4.5}$$

$$= Sq^2(u, Sq^{4k} \hat{\sigma}_{4k+1}) \otimes v^2 \tag{4.6}$$

$$= \phi(u) \otimes v^2 \tag{4.7}$$

(4.2) follows from (2.3), (4.2) and (4.7) from (3.2), (4.4) from the naturality of $Sq^2{}_f$, (4.5) from the Cartan formula, and (4.6) from (2.5). In the case $k = 0$, one needs $v^2 = 0$, in order that $\phi(u \otimes v)$ be defined and (4.5) yields $\phi(u \otimes v) = 0$.

We continue with the proof of (1.6). Let $v_{4l}(N) \in H^{4l}(N)$ be the class such that $z^2 = z v_{4l}(N)$ for all $z \in H^{4l}(N)$. Recall, $v_{4l}(N)^2([N]) = \operatorname{index} N \bmod 2$ (the proof of this is contained in the argument below). Let $u \in H^{4l}(N)$ be a class such that $u = 0$ if $v_{4l}(N) = 0$, $u = v_{4l}(N)$ if $v_{4l}(N)^2 \neq 0$ and $u v_{4l}(N) \neq 0$ if $v_{4l}(N) \neq 0$ and $v_{4l}(N)^2 = 0$. Let $V \subset H^{4l}(N)$ be the subspace spanned by u and $v_{4l}(N)$ and let $U \subset H^{4l}(N)$ be its orthogonal complement, that is, $U = \{z \in H^{4l}(N) \mid zu = z v_{4l}(N) = 0\}$. $H^{4k+1}(M) \otimes U$ is the orthogonal complement of $H^{4k+1}(M) \otimes V$ in $H^{4(k+l)+1}(M \times N)$. Hence a symplectic basis for each of these subspaces will provide a symplectic basis for $H^{4(k+l)+1}(M \times N)$. By (4.1) $H^{4k+1}(M) \otimes U$ makes no contribution to $\psi\{M \times N\}$ as $z^2 = zv_{4l}(N) = 0$ if $z \in U$. Let $\{x_i, y_i\}$ be a symplectic basis for $H^{4k+1}(M)$. We now consider four cases.

Case I. $v_{4l}(N) = 0$. Then $V = 0$ and $\psi\{M \times N\} = 0 = \psi\{M\}v_{4l}{}^2(N)$.

Case II. $v_{4l}{}^2(N) = 0$, $v_{4l}(N) \neq 0$. A symplectic basis for $H^{4k+1}(M) \otimes V$ is given by $\{x_i \otimes v_{4l}(N), y_i \otimes v_{4l}(N)\}$ as the first group of terms and $\{x_i \otimes (v_{4l}(N) + u), y_i \otimes u\}$ as the second group. By (4.1) $\phi(x_i \otimes v_{4l}(N)) = \phi(y_i \otimes v_{4l}(N)) = 0$. Hence $\psi\{M \times N\} = 0$.

Case III. $v_{4l}{}^2(N) \neq 0$, $k > 0$. $\{x_i \otimes v_{4l}(N), y_i \otimes v_{4l}(N)\}$ is a symplectic basis for $H^{4k+1}(M) \otimes V$. Therefore by (4.1),

$$\begin{aligned}\psi\{M \times N\} &= \sum \phi(x_i \otimes v_{4l}(N))([M \times N])\phi(y_i \otimes v_{4l}(N))([M \times N]) \\ &= \sum \phi(x_i)([M])\phi(y_i)([M])v_{4l}{}^2(N) \\ &= \psi(\{M\})I(\{N\}).\end{aligned}$$

Case IV. $v_{4l}{}^2(N) \neq 0$, $k = 0$. The generator of $\Omega_2(SU)$ is represented by $M = S^1 \times S^1$ with the non-trivial SU reduction of its normal bundle. Let $x \in H^1(S^1)$ be the generator. $\{x \otimes 1 \otimes v_{4l}(N), 1 \otimes x \otimes v_{4l}(N)\}$ is a symplectic basis for $H^1(M) \otimes V$. Hence

$$\psi\{M \times N\} = \bar{\phi}\{M \times N, x \otimes 1 \otimes v_{4l}(N)\}\bar{\phi}\{M \times N, 1 \otimes x \otimes v_{4l}(N)\}.$$

By a symmetry argument this equals $\bar{\phi}\{M \times N, x \otimes 1 \otimes v_{4l}(N)\}$. By Wu formulas $v_{4l}(N)$ is a polynomial in the Stiefel-Whitney classes of N. Therefore $v_{4l}(N) = \nu_N{}^*(z_{4l})$ where $z_{4l} \in H^{4l}(BSU)$. z_{4l} is the reduction mod 2 of an integer class $\hat{z}_{4l}$. Hence if $\partial P = S^1 \times N - N'$

$$\partial(P, \nu_p{}^* z_{4l}) = (S^1 \times N, 1 \otimes v_{4l}(N)) - (N', \nu_{N'}{}^* z_{4l})$$

By (3.1) we may choose N' to be 1-connected. Let $y = \nu_{N'}{}^* z_{4l}$. $y^2 \in H^{8l}(N') \approx H_1(N') = 0$. Therefore $\phi(x \otimes y)$ is defined. Let $\hat{y} = \nu_{N'}{}^* \hat{z}_{4l}$ and let $\hat{x} \in H^1(S^1; Z)$ be the generator. By (2.3) and (2.4),

$$\psi\{M \times N\} = \hat{\phi}(\hat{x} \otimes \hat{y})([S^1 \times N'])$$

$$\begin{aligned}\hat{\phi}(\hat{x} \otimes \hat{y}) &= Sq^2{}_{\hat{x} \otimes \hat{y}}(Sq^{4l}\hat{\sigma}_{4l+1}) \\ &= Sq^2{}_{\mathrm{id} \times \hat{y}}(Sq^{4l}(\hat{x} \otimes \hat{\sigma}_{4l})) \\ &= Sq^2{}_{\mathrm{id} \times \hat{y}}(x \otimes \sigma_{4l}{}^2) \\ &= x \otimes Sq^2{}_{\hat{y}}(\sigma_{4l}{}^2).\end{aligned}$$

Hence we must show that $Sq^2{}_{\hat{y}}(\sigma_{4l}{}^2) \neq 0$.

Let ν be the normal bundle of N' embedded in R^{8l+2r} for large r. Let ζ_r be the canonical, real $2r$ bundle over BSU_r. let $\lambda\colon \nu \to \zeta_r$ be the SU reduction of ν, $T(\nu)$ and $T(\zeta_r)$ the Thom spaces, $U \in H^{2r}(T(\zeta_r);Z)$ the Thom class, and let $f\colon S^{8l+2r} \to T(\nu)$ be the map obtained by the Thom construction. Note $T(\lambda)f$ is homotopic to $g\eta$ where η is suspension of the Hopf map and $g\colon S^{8l+2r-1} \to T(\zeta_r)$ is the map obtained from N by the Thom construction. This is because N' and $S^1 \times N$ are SU cobordant.

$$Sq^2{}_{\hat{y}}(\sigma_{4l}{}^2) = Sq^2{}_{\nu_{N'}}(z_{4l}{}^2) \text{ as } \hat{y} = \hat{z}_{4l} \circ \eta_{N'}.$$

Note $Sq^1U = w_1U = 0$ and $Sq^2U = w_2U = 0$.

$$\begin{aligned} f^*(Sq^2{}_{\nu_{N'}}(z_{4l}{}^2) \cdot T(\lambda)^*U) &= f^*Sq^2{}_{T(\lambda)}(z_{4l}{}^2U) \\ &= Sq^2{}_{T(\lambda)f}(z_{4l}{}^2U) \\ &= Sq^2{}_{g\eta}(z_{4l}{}^2U) \\ &= Sq^2{}_{\eta}(g^*z_{4l}{}^2U). \end{aligned} \tag{4.8}$$

But $Sq^2{}_\eta$ is an isomorphism and $g^*(z_{4l}{}^2U)([S^{8k+2r-1}]) = v_{4l}{}^2([N]) \neq 0$. This completes the proof of (1.6).

We next prove (1.1). Suppose $\gamma = \{M\} \in \Omega_{8k+2}(SU)$ where M is stably parallelizable. We must show that $\psi(\gamma) = 0$. Conner and Floyd [7] and Lashof and Rothenberg (unpublished) show that if $\mu \in \Omega_{2n}(SU)$ has all its Chern numbers zero, then $\mu = \alpha\beta$ where $\alpha \in \Omega_2(SU)$ is the generator and $\beta \in \Omega_{2n-2}(SU)$. Hence $\gamma = \alpha\beta$. By (1.6) $\psi(\gamma) = \psi(\alpha)I(\beta) = I(\beta) = v_{4l}(N)^2([N])$ where $v_{4l}(N)$ and N are as above. Suppose $v_{4l}{}^2(N) \neq 0$. Then by (4.3) in Case IV above, $Sq^2{}_{g\eta}(z_{4l}{}^2U) \neq 0$.

Let θ be the secondary cohomology operation associated with the relation $Sq^2Sq^2 = 0$ on integer classes.

$$\theta\colon H^q(X;Z) \cap \operatorname{Ker} Sq^2 \to H^{q+3}(X)/Sq^2H^{q+1}(X).$$

Let θ_f denote the associated functional operation. In [14] it is shown that

$$\theta_{\eta\eta}x = Sq^2{}_\eta Sq^2{}_\eta x. \tag{4.9}$$

$Sq^2(z_{4l}{}^2U) = 0$, hence $\theta(z_{4l}{}^2U) \in H^{8l+3+2r}(T(\zeta_r)) = 0$ is defined and is zero.

$$\begin{aligned} Sq^2{}_\eta Sq^2{}_\eta(z_{4l}{}^2U) &= Sq^2{}_\eta Sq^2{}_\eta(g^*(z_{4l}{}^2U)) \\ &= \theta_{\eta\eta}(g^*(z_{4l}{}^2U)) \\ &= \theta_{g\eta\eta}(z_{4l}{}^2U). \end{aligned}$$

We show that $\theta_{g\eta\eta}(z_{4l}{}^2U)$ is zero and has zero indeterminacy. Since $Sq^2{}_\eta$ is an isomorphism, this shows that $Sq^2{}_{g\eta}(z_{4l}{}^2U) = 0$, which is the contradiction we seek. Note $g\eta\eta\colon S^{8l+2+2r} \to T(\zeta_r)$ is the map corresponding to $S^1 \times S^1 \times N$

under the Thom construction. By hypothesis, $S^1 \times S^1 \times N$ is SU cobordant to a stably parallelizable manifold. Hence $g\eta\eta$ is homotopic to ih where $i\colon S^{2r} \to T(\zeta_r)$ is the inclusion of a fibre and $h\colon S^{8l+2+2r} \to S^{2r}$. Therefore,

$$\theta_{g\eta\eta}(z_{4l}{}^2 U) = \theta_h(i^* z_{4l}{}^2 U) = 0.$$

The indeterminacy of $\theta_{g\eta\eta}(z_{4l}{}^2 U)$ is

$$\begin{aligned}(g\eta\eta)^*(H^{8l+2r}(T(\zeta_r))) &+ \theta(H^{8l+2r-3}(S^{8l+2r})) \\ &+ Sq^2(H^{8l+2r-2}(S^{8l+2r})) + Sq^2{}_{g\eta\eta}(H^{8l+2r-1}(T(\zeta_r))) = 0.\end{aligned}$$

The above argument shows that if $\beta \in \Omega_{8l}(SU)$ and $f\colon S^{8l+2+2r} \to T(\zeta_r)$ is the map associated with $\alpha\beta$, then

$$\psi(\alpha\beta) = \theta_f(z_{4l}{}^2 U)([S^{8l+2+2r}])$$

By examining how the operation θ_f is related to the Thom isomorphism one may prove:

THEOREM 4.10. *If $\alpha \in \Omega_2(SU)$ is the generator and $\beta \in \Omega_{8l}(SU)$, then*

$$\psi(\alpha\beta) = \theta_{\nu_M}(z_{4l}{}^2)([M])$$

where M is a 3-connected manifold representing $\alpha\beta$.

Appendix 1.

We state here, without proof, those parts of our results which go through for $\psi\colon \Omega_{8k+2}(Spin) \to Z_2$.

The first difficulty in generalizing our results to the *Spin* case is that $\bar{\phi}\colon \Omega_{2m}(K(Z_2, m); Spin) \to Z_2$ depends on the choice of the operation ϕ associated to the given relation among primary operations. One choice would be to choose ϕ so that the third suspension of ϕ is zero on all classes of dimension $m - 3$. This is possible and gives rise to two choices for ϕ, each of which give the same $\bar{\phi}$.

The only part of 1.6 that we can prove is the case $k = 0$, namely:

THEOREM A1.1. *If $\alpha \in \Omega_2(Spin)$ is the generator and $\beta \in \Omega_{8k}(Spin)$,*

$$\psi(\alpha\beta) = I(\beta)$$

where $I(\beta)$ is the index of β mod 2.

This theorem follows from the arguments used to prove (1.6), except that one must use slightly more complicated cohomology operations. One also sees that $\psi(\alpha\beta)$ is independent of the choice of ϕ.

Recall, quaternionic projective space PQ_n admits a *Spin* structure and has Euler characteristic 1 if n is even. Hence,

COROLLARY A1.2. *If $\alpha \in \Omega_2(Spin)$ is the generator,*

$$\psi(\alpha\{PQ_{2k}\}) = 1$$

The proof of (1.6) breaks down in the *Spin* case for $k > 0$ because $H^{4(k+l)+1}(M \times N)$ is very complicated even if M and N are simplified by surgery. Also, it is not clear whether an analogous theorem to 4.5 goes through in the spin case.

Appendix 2.

In this appendix we give a proof, due to John Milnor, that a differentiable manifold of the same homotopy type as a Kervaire manifold ([8]) is stably parallelizable.

Let n be odd and $n \neq 1, 3, 7$, let $p: T \to S^n$ be the tangent disc bundle of S^n, let D^n be the closed n-disc, $h: D^n \to S^n$ a homeomorphism into, $k: D^n \times D^n \to T$ a bundle map covering h, $\bar{T}$ a copy of T and let $L = T \cup \bar{T}$ with $h(x, y)$ identified to $h(y, x)$ for each $(x, y) \in D^n \times D^n$. L is a manifold with boundary and ∂L is homeomorphic to S^{2n-1}. Let $f: S^{2n-1} \to \partial L$ be a homeomorphism and let $K^{2n} = L \bigcup_f D^{2n}$. K^{2n} is the manifold constructed by Kervaire.

THEOREM A2.1. *If M is a differentiable manifold with the same homotopy type as K^{2n}, then M is stably parallelizable.*

Proof. Let S^n, $\bar{S}^n$ and $S_1{}^n \subset K^{2n}$ be the zero cross-section of T, the zero cross-section of $\bar{T}$ and a cross-section of the associated sphere bundle of T, respectively. Clearly K^{2n} has a cell structure $S^n \vee \bar{S}^n \cup e^{2n}$ and hence, since S^n and $S_1{}^n$ are isotopic,

$$K^{2n} = S_1{}^n \vee \bar{S}^n \cup e^{2n}. \qquad K^{2n}/S_1{}^n = \bar{S}^n \cup e^{2n}. \qquad K^{2n}/K^{2n} - \operatorname{Int} T$$

is the Thom space of $\tau(S^n)$ which is $S^n \bigcup_{[\iota,\iota]} e^{2n}$. Consider the quotient map

$$u: K^{2n}/S_1{}^n \to K^{2n}/K^{2n} - \operatorname{Int} T$$

u is of degree one on the n and $2n$ cells and hence is a homotopy equivalence. Therefore $K^{2n}/S^n = K^{2n}/S_1{}^n = S^n \bigcup_{[\iota,\iota]} e^{2n}$.

Let $g: K^{2n} \to M$ be a homotopy equivalence. We first show that M is almost parallelizable by showing that $g^*\tau(M)$ is trivial on $S^n \vee \bar{S}^n$. Choose g so that $g \mid S^n$ is a smooth embedding and let ν be the normal bundle of $g(S^n)$. Let $T(\nu)$ be the Thom space of ν, $t: M \to T(\nu)$ the usual map and

$\bar{t}\colon K^{2n} \to K^{2n}/S^n$ the quotient map. $tg \mid S^n$ is homotopically trivial. Hence, up to homotopy, $tg = \bar{g}\bar{t}$ for some $\bar{g}$. One may easily check that $\bar{g}$ is a homotopy equivalence. Hence $T(\nu) = S^n \bigcup_{\alpha} e^{2n}$ where $\alpha = [\iota, \iota]$. Recall, $\alpha = J(\beta)$, $\beta \in \pi_{n-1}(0_n)$ where β is the characteristic class of ν. The stable J homomorphism on $\pi_{n-1}(0)$ is a monomorphism and hence β is stably trivial. Hence ν is stably trivial and therefore $g^*\tau(M) \mid S^n$ is trivial. The same argument shows that $g^*\tau(M) \mid \bar{S}^n$ is trivial.

Recall, the obstruction to an almost parallelizable m-manifold being stably parallelizable is in the kernel of J on $\pi_{m-1}(0)$. J is a monomorphism on $\pi_{2n-1}(0)$. Hence M is stably parallelizable.

BRANDEIS UNIVERSITY.
M. I. T.

REFERENCES.

[1] J. Adem, "Un criterio cohomologico para delerminar composiciones esenciales de transformaciones," *Boletin de la Sociedad Matematica Mexicana*, vol. 1 (1956), pp. 38-48.

[2] C. Arf, "Untersuchungen über quadratische formen in körpern der charakteristik 2," *Crelles Mathematical Journal*, vol. 183 (1941), pp. 148-167.

[3] W. Browder, *Homotopy type of differentiable manifolds*, Colloquium on Algebraic Topology, Aarhus University, 1962, pp. 42-46.

[4] E. H. Brown, Jr., "Note on an invariant of Kervaire," *Michigan Mathematical Journal*, vol. 12 (1965), pp. 23-24.

[5] E. H. Brown, Jr. and F. P. Peterson, "Whitehead products and cohomology operations," *Quarterly Journal of Mathematics*, vol. 15 (1964), pp. 116-120.

[6] E. H. Brown, Jr. and F. P. Peterson, "The Kervaire invariant of $(8k+2)$-manifolds," *Bulletin of hhe American Mathematical Society*, vol. 71 (1965), pp. 190-193.

[7] P. E. Conner and E. E. Floyd, "The *SU*-bordism theory," *Bulletin of the American Mathematical Society*, vol. 70 (1964), pp. 670-675.

[8] M. A. Kervaire, "A manifold which does not admit any differentiable structure," *Commentarii Mathematici Helvetici*, vol 34 (1960), pp. 257-270.

[9] M. A. Kervaire and J. W. Milnor, "Groups of homotopy spheres: I," *Annals of Mathematics*, vol. 77 (1963), pp. 504-537.

[10] R. Lashof, *Some theorems of Browder and Novikov on homotopy equivalent manifolds with an application*, Mimeographed Notes, University of Chicago.

[11] J. W. Milnor, "Remarks concerning spin manifolds," to appear.

[12] S. P. Novikov, "Diffeomorphisms of simply connected manifolds," *Soviet Mathematics, American Mathematical Society*, vol. 3 (1962), pp. 540-543.

[13] F. P. Peterson and N. Stein, "Secondary cohomology operations: two formulas," *American Juornal of Mathematics*, vol. 81 (1959), pp. 281-305.

[14] F. P. Peterson, "Functional cohomology operations," *Transactions of the American Mathematical Society*, vol. 86 (1957), pp. 197-211.

THE MAXIMAL IDEAL SPACE OF A RING OF MEASURABLE FUNCTIONS.

By Hugh Gordon.

The bidual of a sphere of continuous functions, and more generally the dual of a space of measures, are subjects which have been studied for some time. We propose here to show how the concept of a maximal ideal space as applied to the ring of measurable functions is related to that of the dual of a space of measures.

For a moment, let us discuss the bidual of the Banach space $\mathfrak{C}$ of bounded continuous functions of a completely regular space X. Here as elsewhere in the paper, function is understood to mean real-valued function. As proved by Kakutani [5, pp. 1021-1023], the dual of the M-space $\mathfrak{C}$ is an L-space, and the dual of an L-space is again an M-space, in fact, is the space of all continuous functions on a compact Hausdorff space Z. To repeat, the bidual of a space of continuous functions is again a space of continuous functions. However, the relationship between X and Z is very complicated. In fact, to construct Z using Kakutani's ideas, one must first construct the bidual. Then, for example, one can take for Z a suitable subspace of the third dual with its weak*-topology.

An application of the main result of the present paper is an alternative method of constructing a space with the essential property of Z; in fact, Z can be obtained as the Stone-Čech compactification of this other space. More explicitly, given a completely regular space X, we show how to construct a completely regular space Y such that the Banach space of bounded continuous functions on Y is the bidual of that on X. To begin with, we reduce the general case to the one where X is compact by replacing X by its Stone-Čech compactification. Then the dual of the space of continuous functions on X may be regarded as the space of regular Borel measures on X. It turns out that we can easily generalize by replacing X with its Borel structure by an abstract measurable space and replacing the regular measures by a more or less arbitrary space of measures. In the body of the paper then, we consider

---

Received June 2, 1965.
Revised received February 7, 1966.

the more general situation, and we return to the special one only in the last section.

We shall proceed as follows. First we study the maximal ideal space $\hat{X}$ of the space of bounded measurable functions on a measurable space X. Then, a space of measures being given, we choose a suitable subset Y of $\hat{X}$ and give it an appropriate topology. Then using the fact that the dual of L_1 is L_∞ and the Radon-Nikodym theorem, we show that there is a natural correspondence between the bounded continuous functions on Y and the continuous linear functionals on the given space of measures. Finally, we elaborate somewhat on certain special cases, including the one which concerns the bidual of a space of continuous functions.

The reader will recognize without our commenting that the basic material of the second section of the paper is developed in close analogy with material in the well-known book of Gillman and Jerison [4]. The definition of the set X_μ at the beginning of Section 3 is a reduction to the case at hand of a concept introduced in more general circumstances by H. Nakano [8, § 48]. The theorems of Section 4 are essentially special cases of theorems of Nakano. Nevertheless, we have included in Section 4 direct proofs of the results we need later, because these proofs are so short and in order that our treatment may be self-contained. The papers of Kaplan, especially [6], have been helpful to the author in indicating the direction to proceed, although we do not explicitly rely on Kaplan's results at any point. Certainly the reader will be aware that the work of M. H. Stone forms the foundation on which everything involved here rests.

1. Basic definitions. Let $(X, \mathfrak{R})$ be a measurable space. In other words, X is a set and $\mathfrak{R}$ is a σ-ring of subsets of X. We suppose $X \in \mathfrak{R}$ and note that, with this hypothesis, $\mathfrak{R}$ is closed under countable union, countable intersection and complementation. We further suppose that whenever x and y are distinct points of X, there is an $A \in \mathfrak{R}$ with $x \in A$ but $y \notin A$. We shall always understand by measurable set a set belonging to $\mathfrak{R}$.

We make $\mathfrak{R}$ into a commutative ring with unity in the usual way, i. e., by defining $A + B = (A \backslash B) \cup (B \backslash A)$ and $AB = A \cap B$ for all measurable A and B.

We call a real-valued function f on X measurable if $f^{-1}[\alpha, \beta]$ is measurable for every choice of the real numbers $\alpha \leqq \beta$. We denote by $\mathfrak{B}$ the linear space of all bounded measurable functions. We make $\mathfrak{B}$ into a Banach space by defining:

$$\| f \| = \sup\{ | f(x) | \mid x \in X\} \text{ for each } f \in \mathfrak{B}.$$

Note that not only is $\mathfrak{B}$ closed under uniform convergence, but it is even closed under pointwise convergence of sequences. Note further that $\mathfrak{B}$ separates points of X.

Now we construct a compact Hausdorff space $\hat{X}$, of which X is a subset, in one of several well-known ways. By whatever means $\hat{X}$ is constructed, we find it has all of the following properties:

1) Every function $f \in \mathfrak{B}$ has a unique continuous extension to $\hat{X}$. (Henceforth the symbol ^ will always denote continuous extension to $\hat{X}$; thus $\hat{f}$ means that continuous function on $\hat{X}$ which extends $f \in \mathfrak{B}$.)

2) $\mathfrak{B} = C(\hat{X})$, the space of all continuous functions on $\hat{X}$.

3) X is a dense subset of $\hat{X}$.

4) The obvious mapping of $\hat{X}$ into the space of linear functionals on $\mathfrak{B}$ is injective and maps $\hat{X}$ onto the set of non-zero multiplicative linear functionals on $\mathfrak{B}$.

5) The points of $\hat{X}$ are in one-to-one correspondence with the maximal ideals of $\mathfrak{B}$, made into a ring by pointwise multiplication.

As an example of some of the properties and notation just introduced, and because we do need it later, we insert here a lemma.

1.1. LEMMA. *For each* $A \in \mathcal{R}$, $\mathrm{cl}_{\hat{X}}(X \backslash A) = \hat{X} \backslash \mathrm{cl}_{\hat{X}} A$.

Proof. Note first that $\mathrm{cl}(X \backslash A) \cup \mathrm{cl}\, A = \mathrm{cl}\, X = \hat{X}$. Now note that the characteristic function χ_A of A belongs to $\mathfrak{B}$ and hence has a continuous extension $\hat{\chi}_A$ to $\hat{X}$. Thus

$$\mathrm{cl}(X \backslash A) \subset \{p \in \hat{X} \mid \hat{\chi}_A(p) = 0\} \text{ and } \mathrm{cl}\, A \subset \{p \in \hat{X} \mid \hat{\chi}_A(p) = 1\}$$

are disjoint; which completes the proof.

2. $\hat{X}$ as the space of *m*-ultrafilters.

2.1. We call a non-empty subset $\mathcal{F}$ of $\mathcal{R}$ an *m*-filter if:

1) $\emptyset \notin \mathcal{F}$.

2) If $A \in \mathcal{R}$, $B \in \mathcal{F}$ and $A \supset B$, then $A \in \mathcal{F}$.

3) If $A \in \mathcal{F}$ and $B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$.

An *m*-filter which is maximal in the set of all *m*-filters ordered by inclusion will be called an *m*-ultrafilter. The following statement is easily proved and may be found in [3, Chap. 1, § 6, Exercises 17 and 20]: (Note that an *m*-filter is simply a *préfiltre* for the lattice of measurable sets.) If $\mathcal{F}$ is an *m*-

ultrafilter, then $\{X \backslash A \mid A \in \mathfrak{F}\}$ is a maximal ideal of the ring $\mathfrak{R}$; and conversely.

2.2. For each $p \in \hat{X}$ we define $\mathfrak{F}^p$ by $\mathfrak{F}^p = \{A \in \mathfrak{R} \mid p \in \mathrm{cl}_{\hat{X}} A\}$. The following theorem shows that the mapping $p \rightarrow \mathfrak{F}^p$ is a bijection of $\hat{X}$ onto the set of all m-ultrafilters. Towards understanding the theorem, note that any m-ultrafilter may be regarded as a filter base on $\hat{X}$.

THEOREM. *For each $p \in \hat{X}$, $\mathfrak{F}^p$ is an m-ultrafilter, $\mathfrak{F}^p$ converges to p and $\mathfrak{F}^p$ is the only m-ultrafilter which even has p for a point of adherence. Every m-ultrafilter is $\mathfrak{F}^p$ for some $p \in \hat{X}$.*

Proof. Let $p \in \hat{X}$ be given. Clearly $\mathfrak{F}^p$ satisfies 1) and 2) of the definition of an m-filter (2.1). As a consequence of Lemma 1.1 we have $\mathfrak{F}^p = \{A \in \mathfrak{R} \mid p \notin \mathrm{cl}_{\hat{X}}(X \backslash A)\}$. Thus if $A \in \mathfrak{F}^p$ and $B \in \mathfrak{F}^p$,

$$p \notin \mathrm{cl}(X \backslash A) \cup \mathrm{cl}(X \backslash B) = \mathrm{cl}[X \backslash (A \cap B)];$$

which verifies 3) of the definition. Thus $\mathfrak{F}^p$ is an m-filter. Since for all $A \in \mathfrak{R}$ either $p \in \mathrm{cl}\, A$ or $p \in \mathrm{cl}(X \backslash A)$, either $A \in \mathfrak{F}^p$ or $X \backslash A \in \mathfrak{F}^p$. It follows $\mathfrak{F}^p$ is an m-ultrafilter.

Clearly p is a point of adherence of $\mathfrak{F}^p$; in fact, p is a point of adherence of precisely those filter bases on $\hat{X}$ which are contained in $\mathfrak{F}^p$. Thus $\mathfrak{F}^p$ is the only m-ultrafilter with p for a point of adherence.

We recall that in a compact Hausdorff space, like $\hat{X}$, every filter base has at least one point of adherence and every filter base with just one point of adherence converges to that point of adherence. The first remark, together with the preceding paragraph, guarantees that each m-ultrafilter is $\mathfrak{F}^p$ for some p. Thus to complete the proof we need only show that p is the only point of adherence of $\mathfrak{F}^p$. Suppose $q \in \hat{X}$ and $q \neq p$. Then we can find $f \in \mathfrak{B}$ such that $\hat{f}(p) = 0$ but $\hat{f}(q) = 1$. Let $A = \{x \in X \mid f(x) \leqq 1/2\}$; then $\hat{f}(r) \geqq 1/2$ for all $r \in \mathrm{cl}_{\hat{X}}(X \backslash A)$ and hence $p \notin \mathrm{cl}(X \backslash A)$. It follows that $p \in \mathrm{cl}\, A$ and hence $A \in \mathfrak{F}^p$. But $\hat{f}(r) \leqq 1/2$ for all $r \in \mathrm{cl}\, A$; hence $q \notin \mathrm{cl}\, A$. Consequently q is not a point of adherence of $\mathfrak{F}^p$. Thus in fact p is the only point of adherence of $\mathfrak{F}^p$, which completes the proof.

2.3. We have already noted that there is a correspondence between the m-ultrafilters and the maximal ideals of the ring $\mathfrak{R}$; the complements in X of the sets belonging to an object of either kind constitute an object of the other kind. We denote by $\mathfrak{M}^p$ the maximal ideal corresponding to the m-ultrafilter $\mathfrak{F}^p$. As noted during the proof of the theorem, it is a consequence of Lemma 1.1 that for each $A \in \mathfrak{R}$ either $A \in \mathfrak{F}^p$ or $X \backslash A \in \mathfrak{F}^p$. Combining the last two remarks, we have: For any $p \in X$ and $A \in \mathcal{R}$, just one of the following two statements holds:

(i) Both $A \in \mathfrak{F}^p$ and $X \backslash A \in \mathfrak{M}^p$.

(ii) Both $X \backslash A \in \mathfrak{F}^p$ and $A \in \mathfrak{M}^p$.

3. The set X_μ for a fixed measure. By a measure we shall always understand a countably additive real-valued function on $\mathcal{R}$. Note in particular that μX is defined and finite for each measure μ and that μA may be negative. The set of all measures, with the obvious operations, is a linear space. In fact, it can be shown to be a vector lattice if $\mu \geqq \nu$ mean $\mu A \geqq \nu A$ for all measurable sets A. Thus in particular whenever μ and ν are measures, the set $\{\mu, \nu\}$ has a least upper bound in the set of all measures; we denote this l. u. b. by $\mu \vee \nu$. Likewise $\mu \wedge \nu$ denotes the g. l. b. of $\{\mu, \nu\}$. We denote $\mu \vee 0$ by μ^+, $(-\mu) \vee 0$ by μ^- and $\mu^+ + \mu^-$ by $|\mu|$. Note that μ^+, μ^- and $|\mu|$ are positive measures, ν being positive meaning $\nu A \geqq 0$ for all $A \in \mathcal{R}$.

3.1. For each positive measure μ we denote by X_μ the subset of $\hat{X}$ which consists of those $p \in \hat{X}$ for which the m-ultrafilter $\mathfrak{F}^p$ contains no set of μ-measure zero. Equivalently, X_μ consists of those p for which the maximal ideal $\mathfrak{M}^p$ contains every set of μ-measure zero. Since the set of all sets of μ-measure zero is an ideal in $\mathcal{R}$, it is clear that X_μ is never empty unless $\mu = 0$. A description of X_μ in terms of the space $\mathfrak{B}$ of measurable functions is given by the following theorem.

3.2. THEOREM. *X_μ is the set of those $p \in \hat{X}$ with the following property: whenever $f \in \mathfrak{B}$ and $\int |f| \, d\mu = 0$, then $\hat{f}(p) = 0$.*

Proof. First suppose $p \in X_\mu$, $f \in \mathfrak{B}$ and $\int |f| \, d\mu = 0$; we show $f(p) = 0$. Let $\epsilon > 0$ be given; then the set $A = \{x \in X \mid |f(x)| > \epsilon\}$ is of measure zero. Since $p \in X_\mu$, $A \in \mathfrak{M}^p$, and hence $\{x \in X \mid |f(x)| \leqq \epsilon\} = X \backslash A \in \mathfrak{F}^p$. From the definition of $\mathfrak{F}^p$ it follows that $p \in \mathrm{cl}_{\hat{X}}(X \backslash A) \subset \{q \in \hat{X} \mid |\hat{f}(q)| \leqq \epsilon\}$; thus $\hat{f}(p) \leqq \epsilon$. Since this inequality hold for all $\epsilon > 0$, $\hat{f}(p) = 0$.

Now suppose conversely that p has the property; we show $p \in X_\mu$. We must show $A \in \mathfrak{M}^p$, or equivalently $X \backslash A \in \mathfrak{F}^p$, whenever $\mu A = 0$. Let $\chi_A \in \mathfrak{B}$ be the characteristic function of such an A. Then from the property we have $\chi_A(p) = 0$. Since χ_A assumes the value 1 throughout $\mathrm{cl}_{\hat{X}} A$, it follows from Lemma 1.1 that $p \in \mathrm{cl}_{\hat{X}}(X \backslash A)$, and hence by the definition of $\mathfrak{F}^p$ we have $X \backslash A \in \mathfrak{F}^p$.

3.3. COROLLARY. *X_μ is closed in $\hat{X}$, hence compact.*

Proof. X_μ is the intersection of all sets of the form $\{p \in \hat{X} \mid f(\hat{p}) = 0\}$ where $f \in \mathfrak{B}$ and $\int |f| \, d\mu = 0$.

3.4. Theorem. *Let $f \in \mathfrak{B}$. In order that $\int |f| \, d\mu = 0$ it is necessary and sufficient that $f(\hat{p}) = 0$ for all $p \in X_\mu$.*

Proof. The necessity follows from the preceding theorem; we prove the condition is sufficient. To this end, let $f \in \mathfrak{B}$ be such $\int |f| \, d\mu \neq 0$. Then for some $\epsilon > 0$, the set $A = \{x \in X \mid |f(x)| \geqq \epsilon\}$ is of positive μ-measure. Let $\mathcal{N} = \{B \in \mathcal{R} \mid \mu(B \cap A) = 0\}$; then clearly $\mathcal{N}$ is a proper ideal in $\mathcal{R}$. Let $\mathcal{M}^p$ be any maximal ideal of $\mathcal{R}$ containing $\mathcal{N}$. Now since $X \backslash A \in \mathcal{N} \subset \mathcal{M}^p$, $A \in \mathfrak{J}^p$, and hence $p \in \mathrm{cl}_{\hat{X}} A \subset \{q \in \hat{X} \mid |\hat{f}(q)| \geqq \epsilon\}$. Since it follows $\hat{f}(p) \neq 0$, it remains only to show $p \in X_\mu$. But certainly each $B \in \mathcal{R}$ of measure zero belongs to $\mathcal{N}$, hence to $\mathcal{M}^p$. That $p \in X_\mu$ now follows from the definition of X_μ.

4. The sets X_μ for varying μ.

4.1. Theorem. *Let μ and ν be positive measures. Then ν is absolutely continuous with respect to μ if and only if $X_\nu \subset X_\mu$.*

Proof. "Only if" follows directly from the definitions; "if" follows from the preceding theorem.

4.2. Theorem. *Let μ and ν be positive measures. Then μ and ν are singular if and only if $X_\mu \cap X_\nu = \emptyset$.*

Proof. First suppose $X_\mu \cap X_\nu = \emptyset$. Since $\mu \wedge \nu$ is absolutely continuous with respect to both μ and ν, $X_{\mu \wedge \nu} \subset X_\mu \cap X_\nu = \emptyset$; hence by the remark in 3.1, $\mu \wedge \nu = 0$.

Conversely suppose $\mu \wedge \nu = 0$. Then there are A and $B \in \mathcal{R}$ such that $B = X \backslash A$ and $\mu A = \nu B = 0$. Then for a hypothetical $p \in X_\mu \cap X_\nu$ we have $A \in \mathcal{M}^p$ and $B \in \mathcal{M}^p$. Therefore, since $\mathcal{M}^p$ is an ideal of the ring $\mathcal{R}$, $A + B = (A \backslash B) \cup (B \backslash A) = X$ belongs to $\mathcal{M}^p$. But X is the unity element of $\mathcal{R}$ and $\mathcal{M}^p$ is a maximal ideal; thus $X_\mu \cap X_\nu = \emptyset$.

4.3 Theorem. *Let μ and ν be positive measures. Then $X_{\mu+\nu} = X_\mu \cup X_\nu$.*

Proof. Since μ and ν are both absolutely continuous with respect to $\mu + \nu$ we have $X_{\mu+\nu} \supset X_\mu \cup X_\nu$. We complete the proof by supposing $p \notin X_\mu \cup X_\nu$ and showing that then $p \notin X_{\mu+\nu}$. Since $p \notin X_\mu$, there is an $A \in \mathfrak{J}^p$ with $\mu A = 0$. Similarly since $p \notin X_\nu$, there is a $B \in \mathfrak{J}^p$ with $\nu B = 0$. We have $A \cap B \in \mathfrak{J}^p$ and $(\mu + \nu)(A \cap B) = 0$; thus $p \notin X_{\mu+\nu}$.

4.4. COROLLARY. *Let μ and ν be positive measures. Then* $X_{\mu\wedge\nu} = X_\mu \cap X_\nu$.

Proof. Let $\mu' = \mu - \mu \wedge \nu$ and $\nu' = \nu - \mu \wedge \nu$; then $\mu' \wedge \nu' = 0$. We have

$$\begin{aligned} X_\mu \cap X_\nu &= (X_{\mu\wedge\nu} \cup X_{\mu'}) \cap (X_{\mu\wedge\nu} \cup X_{\nu'}) \\ &= X_{\mu\wedge\nu} \cup (X_{\mu'} \cap X_{\nu'}) = X_{\mu\wedge\nu} \cup \emptyset = X_{\mu\wedge\nu}. \end{aligned}$$

4.5. THEOREM. *For any positive measures μ and ν, $X_\mu \cap X_\nu$ is both closed and open in X_μ.*

Proof. We may write $\mu = \mu_1 + \mu_2$ where μ_1 is a positive measure absolutely continuous with respect to ν and μ_2 is one singular with respect to ν. By the preceding theorem, $X_{\mu_1} \cup X_{\mu_2} = X_\mu$ and hence $(X_{\mu_1} \backslash X_\nu) \cup (X_{\mu_2} \backslash X_\nu) = X_\mu \backslash X_\nu$. But by Theorem 4.1, $X_{\mu_1} \subset X_\nu$ and by Theorem 4.2, $X_{\mu_2} \cap X_\nu = \emptyset$. Thus we have $X_{\mu_2} = X_\mu \backslash X_\nu$. It follows, since X_{μ_2} and X_ν are each closed in $\hat{\mathfrak{X}}$ (Corollary 3.3), that $X_\mu \backslash X_\nu$ and $X_\mu \cap X_\nu$ are each closed in X_μ. Consequently, they are also open in X_μ.

5. The space Y. Let $\mathfrak{M}$ be a linear subspace of the space of all measures. We suppose that $\mathfrak{M}$ has the following property: Whenever $\mu \in \mathfrak{M}$ and ν is absolutely continuous with respect to $\mathfrak{M}$, then $\nu \in \mathfrak{M}$. More explicitly, if $\mu \in \mathfrak{M}$ and $|\nu| A = 0$ whenever $|\mu| A = 0$, then $\nu \in \mathfrak{M}$. It follows that whenever $\mu \in \mathfrak{M}$, then $\mu^+ \in \mathfrak{M}$ since $0 \leqq \mu^+ \leqq |\mu|$ and $|\mu^+| = \mu^+$. Since $\mu \vee \nu = \nu + (\mu - \nu)^+$ and $\mu \wedge \nu = \mu - (\mu - \nu)^+$; whenever μ and $\nu \in \mathfrak{M}$, then $\mu \vee \nu$ and $\mu \wedge \nu \in \mathfrak{M}$. Thus $\mathfrak{M}$ is a sublattice of the space of all measures.

5.1. Next we define a certain topological space Y. As a set, Y will be a subset of $\hat{\mathfrak{X}}$, but in general it will not be a subspace. (A necessary and sufficient condition that Y be a topological subspace of $\hat{\mathfrak{X}}$ is given by Theorem 7.5.) For this reason, we shall henceforth call the topology of $\hat{\mathfrak{X}}$ used above the β-topology and call the topology of Y about to be defined the δ-topology. The set Y is defined by $Y = \cup \{X_\mu \mid \mu \in \mathfrak{M}, \mu \geqq 0\}$. The δ-topology has for a base (for open sets) the set of all X_μ with $\mu \in \mathfrak{M}$ positive; this definition is permissible by Corollary 4.4.

5.2. THEOREM. *The relative topology induced on Y by the β-topology is coarser than the δ-topology.*

Proof. Let U be open for the relative β-topology on Y and let $p \in U$. To prove the theorem, it suffices to find a positive measure $\nu \in \mathfrak{M}$ such that $p \in X_\nu \subset U$. Henceforth in this proof we use only the β-topology; the δ-topology plays no further role.

We begin by choosing U' open in $\hat{X}$ such that $U' \cap Y = U$. Then working with the space $\hat{X}$, we choose a closed neighborhood V of p contained in the open neighborhood U'. Next we choose a continuous function $\hat{g}$ mapping $\hat{X}$ into $[0,1]$ such that $\hat{g}(p) = 1$ but $\hat{g}(q) = 0$ for all $q \in \hat{X} \backslash V$; let $g = \hat{g} \mid X$ and note $g \in \mathfrak{B}$. Finally, since $p \in Y = \cup \{X_\mu \mid \mu \in \mathfrak{M}, \mu \geqq 0\}$, we may choose a positive measure $\mu \in \mathfrak{M}$ such that $p \in X_\mu$. From the Radon-Nikodym theorem it follows that there is a positive measure ν, absolutely continuous with respect to μ and hence in $\mathfrak{M}$, such that:

$$\int f \, d\nu = \int fg \, d\mu \text{ for all } f \in \mathfrak{B}.$$

We next check that $p \in X_\nu$. For each $f \in \mathfrak{B}$ with $\hat{f}(p) \neq 0$, we have $(fg)\hat{}\,(p) \neq 0$, since $\hat{g}(p) = 1$. Hence for all such f we have

$$\int |f| \, d\nu = \int |f| \, g \, d\mu = \int |fg| \, d\mu \neq 0$$

since $p \in X_\mu$ (Theorem 3.2). Thus also by Theorem 3.2, $p \in X_\nu$.

It remains only to show $X_\nu \subset U$. Suppose on the contrary $q \in X_\nu$ but $q \notin U$. Then since $V \subset U'$, $q \in X_\nu \subset Y$ and $q \notin U = U' \cap Y$; we have $q \notin V$. Thus we can find a continuous function f mapping $\hat{X}$ into $[0,1]$ such that $\hat{f}(q) = 1$ but $\hat{f}(q') = 0$ for all $q' \in V$; set $f = \hat{f} \mid X$. Then $\hat{f}\hat{g}$ vanishes throughout $V \cup (\hat{X} \backslash V) = \hat{X}$; hence $\int |f| \, d\nu = \int |fg| \, d\mu = 0$. But then since $\hat{f}(q) = 1 \neq 0$, we have by Theorem 3.2 $q \notin X_\nu$. This contradiction completes the proof.

5.3. Corollary. *For each positive measure μ, the δ-topology and the β-topology both induce the same relative topology on X_μ.*

Proof. By the theorem the topology induced by the δ-topology is finer than the one induced by the β-topology. But by Theorem 4.5 and Definition 5.1, the former of these induced topologies has a base consisting of sets open for the latter of them.

5.4. Corollary. *Y with its δ-topology is a locally compact Hausdorff space.*

Proof. First note that whenever $p \in X_\mu$, X_μ is a compact neighborhood of p (Corollary 3.3). Since the δ-topology is finer than the topology induced on Y by the β-topology, which is Hausdorff, the δ-topology is Hausdorff.

5.5. Corollary. *Let $A \subset Y$. In order that there be a positive $\mu \in \mathfrak{M}$ with $A = X_\mu$ it is necessary and sufficient that A be open and compact for the δ-topology.*

Proof. Necessity has essentially been proved (Definition 5.1, Corollaries 3.3 and 5.3); we prove sufficiency here. Since A is open, it is union of those X_μ it contains. Since every X_μ is open and A is compact, we can find $\mu_1, \cdots, \mu_n \in \mathfrak{M}$ such that $A = X_{\mu_1} \cup \cdots \cup X_{\mu_n}$. Setting $\mu = \mu_1 + \cdots + \mu_n$, we have as a consequence of Theorem 4.3, $A = X_\mu$.

5.6. COROLLARY. *In order that a mapping ϕ of Y into a topological space Z be continuous for the δ-topology, it is necessary and sufficient that for each positive $\mu \in \mathfrak{M}$, $\phi \mid X_\mu$ be continuous for the relative β-topology.*

Proof. Necessity is a trivial consequence of Corollary 5.3. Sufficiency is also clear using the same corollary, once it is noted that every $p \in Y$ has a δ-neighborhood of the form X_μ.

6. The isomorphism between $\mathfrak{M}^*$ and $C^*(Y)$. Now we introduced the obvious norm on the linear space $\mathfrak{M}$ by defining $\| \mu \| = | \mu | X$ for each $\mu \in \mathfrak{M}$ The purpose of 6.1-6.3 is to show how each bounded real-valued function Φ on Y continuous for the δ-topology gives rise to a bounded linear functional F on $\mathfrak{M}$.

6.1. We define $F\mu$, for $\mu \in \mathfrak{M}$ positive, as follows: The restriction of Φ to X_μ is continuous for the β-topology (Corollary 5.6) and X_μ is compact for this topopolgy (Corollary 3.3). Since every continuous real-valued function on a compact subspace of a completely regular space admits a continuous extension to the entire space, we may choose some $\hat{f}$ continuous on $\hat{X}$ for the β-topology and such that $\hat{f} \mid X_\mu = \Phi \mid X_\mu$. We define $F\mu = \int f\, d\mu$, where of course $f = \hat{f} \mid X \in \mathfrak{B}$. We next show $F\mu$ is independent of the choice of f; suppose f_1 and f_2 are each β-continuous extensions of $\Phi \mid X_\mu$ to $\hat{X}$. Then $\hat{f}_1 - \hat{f}_2 = (f_1 - f_2)^\wedge$ vanishes throughout X_μ, hence by Theorem 3.4, $\int | f_1 - f_2 | \, d\mu = 0$. It follows that $\int f_1\, d\mu = \int f_2\, d\mu$.

6.2. LEMMA. *If μ and $\nu \in \mathfrak{M}$ are positive, $F(\mu + \nu) = F\mu + F\nu$.*

Proof. As above, we may find $f \in \mathfrak{B}$ such that $\hat{f} \mid X_{\mu+\nu} = \Phi \mid X_{\mu+\nu}$. Then since $X_\mu \cup X_\nu = X_{\mu+\nu}$ (Theorem 4.3), $\hat{f} \mid X_\mu = \Phi \mid X_\mu$ and $\hat{f} \mid X_\nu = \Phi \mid X_\nu$. Since then by definition

$$F(\mu + \nu) = \int f\, d(\mu + \nu), \; F\mu = \int f\, d\mu \text{ and } F\nu = \int f\, d\nu,$$

the lemma follows.

6.3. Using the lemma and noting $F(\alpha\mu) = \alpha F\mu$ for all positive α and μ, we may extend F to all of $\mathfrak{M}$ obtaining a linear functional on $\mathfrak{M}$.

We show finally that F is bounded. Choose α such that $\Phi(p) \leqq \alpha$ for all $p \in Y$.

We first consider some positive $\mu \in \mathfrak{M}$ and choose the β-continuous $\hat{f}$ of the definition of $F\mu$ in such a way that $|\hat{f}(p)| \leqq \alpha$ for all $p \in \hat{X}$ as well as $\hat{f} \mid X_\mu = \Phi \mid X_\mu$. Then we have $|F\mu| = |\int f\, d\mu| \leqq \alpha \cdot \mu X$. Using this inequality, we have for arbitrary μ,

$$\begin{aligned} |F\mu| = |F\mu^+ - F\mu^-| &\leqq |F\mu^+| + |F\mu^-| \\ &\leqq \alpha \cdot \mu^+ X + \alpha \cdot \mu^- X = \alpha \cdot |\mu| X = \alpha \|\mu\|. \end{aligned}$$

Thus F is bounded; in fact $\|F\| \leqq \alpha$ whenever $|\Phi| \leqq \alpha$.

6.4. Next we consider the reverse problem of defining Φ when F is given. Let F be a bounded linear functional on $\mathfrak{M}$. Consider first some fixed positive $\mu \in \mathfrak{M}$ and let $\mathfrak{M}_\mu$ denote the space of measures absolutely continuous with respect to μ. Note that by hypothesis $\mathfrak{M}_\mu \subset \mathfrak{M}$. The space $\mathfrak{M}_\mu$ is isomorphic to $L_1(\mu)$, the measure $\nu \in \mathfrak{M}_\mu$ corresponding to its Radon-Nikodym derivative $d\nu/d\mu \in L_1(\mu)$. Thus the restriction of F to $\mathfrak{M}_\mu$ corresponds to an element of $L_\infty(\mu)$. This element of $L_\infty(\mu)$ is an equivalence class of functions in $\mathfrak{B}$. More precisely, it is that subset of $\mathfrak{B}$ which consists of all $f \in \mathfrak{B}$ for which:

$$F\nu = \int f \frac{d\nu}{d\mu} d\mu = \int f\, d\nu \text{ for all } \nu \in \mathfrak{M}_\mu.$$

Note that if f_1 and f_2 each have this property, then f_1 and f_2 differ only on a set of μ-measure zero, and hence by Theorem 3.2 we have $\hat{f}_1 \mid X_\mu = \hat{f}_2 \mid X_\mu$. Thus we may make the following definition:

6.5. *Definition.* Φ_μ is that function on X_μ with the property that $\hat{f} \mid X_\mu = \Phi_\mu$ whenever $f \in \mathfrak{B}$ is such that:

$$F\nu = \int f\, d\nu \text{ for all } \nu \in \mathfrak{M}_\mu.$$

6.6. LEMMA. *Suppose g is any function in $\mathfrak{B}$ such that $\hat{g} \mid X_\mu = \Phi_\mu$. Then $F\nu = \int g\, d\nu$ for all $\nu \in \mathfrak{M}_\mu$.*

Proof. Using the definition of Φ_μ we choose $f \in \mathfrak{B}$ such that $\hat{f} \mid X_\mu = \Phi_\mu = \hat{g} \mid X_\mu$ and $F\nu = \int f\, d\nu$ for all $\nu \in \mathfrak{M}_\mu$. By Theorem 3.4, f and g differ only on a set of μ-measure zero. Since each $\nu \in \mathfrak{M}_\mu$ is absolutely continuous with respect to μ, we have for each such ν, $\int g\, d\nu = \int f\, d\nu = F\nu$.

6.7. Now we permit μ to vary but still keep F fixed.

LEMMA. *Suppose μ_1 and $\mu_2 \in \mathfrak{M}$ are positive and μ_1 is absolutely continuous with respect to μ_2. Then $\Phi_{\mu_1} = \Phi_{\mu_2} \mid X_{\mu_1}$.*

Proof. Let $f \in \mathfrak{B}$ be such that $f \mid X_{\mu_2} = \Phi_{\mu_2}$. Then by the preceding lemma $F\nu = \int f\, d\nu$ for all $\nu \in \mathfrak{M}_{\mu_2}$. Since $\mathfrak{M}_{\mu_1} \subset \mathfrak{M}_{\mu_2}$, we have $\hat{f} \mid X_{\mu_1} = \Phi_{\mu_1}$ by Definition 6.5. Thus, using Theorem 4.1,

$$\Phi_{\mu_1} = \hat{f} \mid X_{\mu_1} = (\hat{f} \mid X_{\mu_2}) \mid X_{\mu_1} = \Phi_{\mu_2} \mid X_{\mu_1}.$$

6.8. Lemma. *Let μ and $\mu' \in \mathfrak{M}$ be arbitrary but positive. Then*

$$\Phi_\mu \mid X_\mu \cap X_{\mu'} = \Phi_{\mu'} \mid X_\mu \cap X_{\mu'}.$$

Proof. We write $\mu = \mu_1 + \mu_2$ with μ_1 absolutely continuous with respect to μ' (as well as with respect to μ) and μ_2 singular with respect to μ'. First note that by the previous lemma, $\Phi_\mu \mid X_{\mu_1} = \Phi_{\mu'} \mid X_{\mu_1}$. Hence it suffices to show $X_{\mu_1} = X_\mu \cap X_{\mu'}$. But since $X_\mu = X_{\mu_1} \cup X_{\mu_2}$ by Theorem 4.3, we have $X_\mu \cap X_{\mu'} = (X_{\mu_1} \cap X_{\mu'}) \cup (X_{\mu_2} \cap X_{\mu'}) = X_{\mu_1} \cup \emptyset = X_{\mu_1}$ by Theorems 4.1 and 4.2.

6.9. Now we are in a position to define Φ corresponding to F. Recall $Y = \cup \{X_\mu \mid \mu \in \mathfrak{M}, \mu \geqq 0\}$. Whenever $p \in Y$, we define $\Phi(p) = \Phi_\mu(p)$ where μ is any measure with $p \in X_\mu$; by the last lemma Φ is well-defined. To show Φ is continuous for the δ-topology it suffices to show the restriction of Φ to each X_μ is β-continuous (Corollary 5.6). But for each μ we can find $f \in \mathfrak{B}$ with $\Phi \mid X_\mu = \Phi_\mu = \hat{f} \mid X_\mu$ and every $\hat{f}$ is β-continuous. Finally we show Φ is bounded by noting that each function, except the derivative, which is mentioned in the preceding discussion can be chosen to be bounded, in fact to be in absolute value $\leqq \| F \|$.

To further specify the relationship between the given F and the Φ just defined, we record that, for each positive measure μ, since $\Phi \mid X_\mu = \Phi_\mu$, we have by Lemma 6.6.

$$F\mu = \int f\, d\mu \text{ whenever } f \in \mathfrak{B} \text{ and } \hat{f} \mid X_\mu = \Phi \mid X_\mu.$$

6.10. Let $C^*(Y)$ denote the space of all bounded continuous real-valued functions on Y and let $\mathfrak{M}^*$ denote the usual dual of $\mathfrak{M}$, i. e. the space of all bounded linear functional on $\mathfrak{M}$. We put the customary norms on $C^*(Y)$ and $\mathfrak{M}^*$. 6.1-6.3 may be interpreted as defining a mapping I of $C^*(Y)$ into $\mathfrak{M}^*$; $I\Phi$ for $\Phi \in C^*(Y)$ is uniquely determined by the statement:

$$I(\Phi)\mu = \int f\, d\mu \text{ whenever } \mu \geqq 0,\ f \in \mathfrak{B} \text{ and } \hat{f} \mid X_\mu = \Phi \mid X_\mu.$$

The final remark of 6.9 implies that I is surjective. Suppose Φ_1 and $\Phi_2 \in C^*(Y)$ and let $\mu \in \mathfrak{M}$ be positive. Then we may choose f_1 and $f_2 \in \mathfrak{B}$ with $\hat{f}_i \mid X_\mu = \Phi_i \mid X_\mu$ $(i = 1, 2)$. Since $(f_1 + f_2)^\wedge \mid X_\mu = (\Phi_1 + \Phi_2) \mid X_\mu$, we have $I(\Phi_1+\Phi_2)\mu=\int(f_1+f_2)\,d\mu=\int f_1\,d\mu+\int f_2\,d\mu=I(\Phi_1)\mu+I(\Phi_2)\mu$. It follows without difficulty that I is linear.

We next show I is injective by showing that $\Phi\neq 0$ implies $I(\Phi)\neq 0$. Replacing Φ by $-\Phi$ if necessary, we suppose there is a $p\in Y$ with $\Phi(p)>0$. Since $\{q\in Y\mid \Phi(q)>0\}$ is open in Y, it is the union of those X_μ it contains (5.1). Thus we may choose a positive $\mu\in\mathfrak{M}$ with $p\in X_\mu$ and $\Phi(q)>0$ for all $q\in X_\mu$. Then we may choose a positive $f\in\mathfrak{B}$ such that $\hat{f}\mid X_\mu=\Phi\mid X_\mu$. We have $I(\Phi)\mu=\int f\,d\mu=\int|f|\,d\mu\neq 0$ by Theorem 3.4. Hence $I(\Phi)\neq 0$; which completes the proof that I is injective.

6.11. To summarize, I is a linear isomorphism of $C^*(Y)$ onto $\mathfrak{M}^*$. Given $\Phi\in C^*(Y)$, the construction of 6.1-6.3 determines $I(\Phi)$; in particular we may conclude that $\|I(\Phi)\|\leqq\|\Phi\|$. It is easy to check that $\Phi\geqq 0$ implies $I(\Phi)\geqq 0$. On the other hand, given $F\in\mathfrak{M}^*$, the construction of 6.4-6.9 determines $I^{-1}(F)$; in particular $\|I^{-1}(F)\|\leqq\|F\|$. Again it is easy to check that $F\geqq 0$ implies $I^{-1}(F)\geqq 0$. Combining all these remarks, have proved the following theorem:

THEOREM. *The mapping I defined by the construction of* 6.1-6.3 *and whose inverse is defined by the construction of* 6.4-6.9 *is an isometry, a linear isomorphism and an order isomorphism of $C^*(Y)$ onto $\mathfrak{M}^*$.*

7. Examples and comments.

7.1. We begin this section by considering the consequences of the simplifying hypothesis that $\mathfrak{M}$ consist of all measures on $\mathcal{R}$. First, with this hypothesis and $x\in X$ arbitrary, the measure μ defined by $\mu A=1$ if $A\in\mathfrak{F}^x$ and $\mu A=0$ if $A\notin\mathfrak{F}^x$ belongs to $\mathfrak{M}$. Since for the m-ultrafilter $\mathfrak{F}^x$ we have $\mathfrak{F}^x=\{A\in\mathcal{R}\mid x\in \mathrm{cl}_{\hat{X}}A\}=\{A\in\mathcal{R}\mid x\in A\}$ (Lemma 1.1), it follows from the definition of X_μ that $X_\mu=\{x\}$. Hence each $x\in X$ is an isolated point of Y, for the δ-topology, which has the sets X_μ for a base.

7.2. We next answer the converse question of whether it is always true even with the present hypothesis, that every isolated point of Y belongs to X. The negative answer is established by the following counterexample. Let X be an uncountable set and let $\mathcal{R}$ consist of all $A\subset X$ such that one of A and $X\backslash A$ is countable. In accordance with our general hypothesis, let $\mathfrak{M}$ consist of all measures. Clearly the set of all those subsets A of X for which $X\backslash A$ is countable is an m-ultrafilter, which we denote by $\mathfrak{F}^p$. As before, let μ be the measure such that $\mu A=1$ if $A\in\mathfrak{F}^p$ and $\mu A=0$ if $A\notin\mathfrak{F}^p$. Again $X_\mu=\{p\}$ and p is an isolated point of Y. But this time $p\notin X$. (There is an important special case in which X is precisely the set of isolated points of Y; this special case is treated in the next section.)

7.3. (Added later) The question of whether, still with the hypothesis that $\mathfrak{M}$ consists of all measures, Y can be a proper subset of $\hat{X}$ has an affirmative answer, even in the special case considered in the next section. This point will be covered in a later paper.

7.4. One consequence of the statement that every point of X is an isolated point of Y is simply the inclusion $X \subset Y$. Since X is β-dense in $\hat{X}$, it follows that X is β-dense in Y. Thus every set which is β-open in Y must meet X. In particular, if some X_μ is β-open in Y, we can choose $x \in X \cap X_\mu$ and conclude that $\mu A > 0$ for each $A \in \mathfrak{F}^x$; i.e. for each $A \in \mathfrak{R}$ with $x \in A$. Thus whenever there is a measure μ for which no such point x exists—and this is usually the case—there is an X_μ which is not β-open in Y. Hence, since each X_μ is always δ-open in Y, the δ-topology on Y often differs from the relative β-topology.

7.5. It is of some interest to drop the hypothesis that $\mathfrak{M}$ consist of all measures and to treat in the general case the question of when the two topologies just mentioned coincide. The following theorem shows how hard it is for them to be identical, but the corollary gives a significant special case in which they are identical.

THEOREM. *In order that the δ-topology of Y be the relative β-topology it is necessary and sufficient that for each positive $\mu \in \mathfrak{M}$ there be an $A \in \mathfrak{R}$ with $\mu A = 0$ such that each positive $\nu \in \mathfrak{M}$ with $\nu A = 0$ is absolutely continuous with respect to μ.*

Proof. Before starting, we note that since the δ-topology is always finer than the relative β-topology (Theorem 5.2), and since the sets X_μ for positive $\mu \in \mathfrak{M}$ form a base for the δ-topology, the two topologies coincide exactly when each X_μ is β-open in Y. In other words, the topologies coincide exactly when for each positive $\mu \in \mathfrak{M}$ there is a U, β-open in $\hat{X}$, with $U \cap Y = X_\mu$. The proof now proceeds without further reference to the δ-topology.

Necessity: Let $\mu \in \mathfrak{M}$ be positive. As just noted, there is a U, open in $\hat{X}$, for which $U \cap Y = X_\mu$. Thus for each $p \in X_\mu$, since U is a neighborhood of p and $\mathfrak{F}^p$ regarded as a filter base on $\hat{X}$ converges to p (Theorem 2.2), we can choose a $B_p \in \mathfrak{F}^p$ with $\mathrm{cl}_{\hat{X}} B_p \subset U$. From the definition of $\mathfrak{F}^p$ it follows that $p \in \mathrm{cl}\, B_p \subset U$. Since for each $p \in X_\mu$, $\mathrm{cl}\, B_p = \hat{X} \backslash \mathrm{cl}\, (X \backslash B_p)$ (Lemma 1.1.) is open and since X_μ is compact, there are $p_1, \cdots, p_n \in X_\mu$ such that

$X_\mu \subset \bigcup\limits_{i=1}^{n} \operatorname{cl} B_{p_n}$. Now we set $B = \bigcup\limits_{i=1}^{n} B_{p_n}$; then $B \in \mathfrak{R}$ and we have $X_\mu \subset \operatorname{cl} B \subset U$ and hence $Y \cap \operatorname{cl} B = X_\mu$. Let $A = X \backslash B$; applying Theorem 3.4 to the characteristic function of A, we see $\mu A = 0$. Now suppose $\nu \in \mathfrak{M}$ is given with $\nu A = 0$. Again considering the characteristic function of A and using Theorem 3.4, we see $X_\nu \cap \operatorname{cl} A = \emptyset$. Thus we have using Lemma 1.1, $X_\nu = Y \cap X_\nu \subset Y \cap (\hat{X} \backslash \operatorname{cl} A) = Y \cap \operatorname{cl} B = X_\mu$; thus by Theorem 4.1, ν is absolutely continuous with respect to μ. In short, we have found A with the required properties.

Sufficiency: Let $\mu \in \mathfrak{M}$; let A be as described in the statement of the theorem. Defining $U = \operatorname{cl} B$ with $B = X \backslash A$, we have $U = \hat{X} \backslash \operatorname{cl} A$ is open in $\hat{X}$; it remains only to show $U \cap Y = X_\mu$. Since $\mu A = 0$, repeating the argument using Theorem 3.4, $X_\mu \cap \operatorname{cl} A = \emptyset$; hence $X_\mu \subset \operatorname{cl} B = U$. Thus $X_\mu \subset U \cap Y$. Towards proving the opposite inclusion, choose any $p \in Y$; there is a positive measure $\nu' \in \mathfrak{M}$ with $p \in X_{\nu'}$. We define a measure ν by $\nu C = \nu'(C \cap B)$ for all $C \in \mathfrak{R}$; then since ν is absolutely continuous with respect to ν', $\nu \in \mathfrak{M}$. Clearly $\nu A = \nu' \emptyset = 0$. Hence, using the condition of the theorem, ν is absolutely continuous with respect to μ, and consequently $X_\nu \subset X_\mu$. Now if also $p \in U \subset \operatorname{cl} B$, $B \in \mathfrak{F}^p$. Hence for each $C \in \mathfrak{F}^p$, we have $B \cap C \in \mathfrak{F}^p$ and thus $\nu C = \nu'(B \cap C) \neq 0$. In other words, $p \in X_\nu$. Combining our remarks, we have for each $p \in U \cap Y$, $p \in X_\nu \subset X_\mu$; i.e. $U \cap Y \subset X_\mu$. Since $U \cap Y \supset X_\mu$ was verified earlier, we have $U \cap Y = X_\mu$.

7.6. Corollary. *Suppose for each $x \in X$, $\{x\} \in \mathfrak{R}$. Suppose also that for each $\mu \in \mathfrak{M}$ and each $A \in \mathfrak{R}$, $\mu A = \sum\limits_{x \in A} \mu\{x\}$. Then the δ-topology of Y coincides with the relative β-topology.*

Proof. We verify the condition of the theorem; let a positive $\mu \in \mathfrak{M}$ be given. Let $B = \{x \in X \mid \mu\{x\} > 0\}$; clearly B is countable and hence belongs to $\mathfrak{R}$. Now if we let $A = X \backslash B$, it is clear A has the required property.

8. The bidual of a space of continuous functions. We now consider the special case already alluded to in the introduction. Let X be a compact Hausdorff space and $\mathfrak{C}$ be the space of continuous functions on X. The space of all continuous linear functionals on $\mathfrak{C}$ is often regarded as being the space of regular Borel measures on X. Although the result of Section six does apply when $\mathfrak{M}$ is this space of measures, we prefer here to regard the dual of $\mathfrak{C}$ as the space of Baire measures on X. More explicitly, let $\mathfrak{B}$ be the space of bounded Baire functions on X; in other words, $\mathfrak{B}$ is the smallest set of bounded functions of X which contains $\mathfrak{C}$ and is closed in the set of all bounded functions under the operation of taking pointwise limits of sequences.

Let $\mathfrak{R}$ consist of those subsets of X whose characteristic functions belong to $\mathfrak{B}$; we call the sets in $\mathfrak{R}$ Baire sets. Then it is not hard to check that X, $\mathfrak{R}$ and $\mathfrak{B}$ satisfy all the hypotheses which we imposed at the beginning of Section 1.

8.1. We let $\mathfrak{M}$ be the space of all measures, in the sense specified in Section 3; i.e. $\mathfrak{M}$ is the space of Baire measures. Then $\mathfrak{M}$ is known to be the dual of $\mathfrak{C}$; consequently we may use 6.11 to identify the bidual of $\mathfrak{C}$ with space $C^*(Y)$ of all bounded continuous functions on our Y. Now consider the canonical embeddings of $\mathfrak{C}$ in its bidual $C^*(Y)$. To describe this embedding explicitly, the value of a function $f \in \mathfrak{C}$, regarded as an element of the bidual, at a measure μ is $\int f\, d\mu$. Let us compare this embedding with the one suggested by our theory. Given any $f \in \mathfrak{C}$, or more generally given any $f \in \mathfrak{B}$, we can extend f to $\hat{X}$ by β-continuity obtaining $\hat{f}$. Then we may consider $\hat{f} \mid Y$, which is δ-continuous, since the δ-topology is finer than the β-topology on Y. Now $\hat{f} \mid Y$ is an element of the bidual of $\mathfrak{C}$. According to the construction of 6.1, the value of this element at the measure μ is $\int f\, d\mu$, precisely as before.

8.2. We next show that X, as a set, is the set of isolated points of Y, of course for the δ-topology. Suppose p is an isolated point of Y. Then since $\{p\}$ is open and compact in Y, there is a positive measure $\mu \in \mathfrak{M}$ with $X_\mu = \{p\}$ (Corollary 5.5). Each f which belongs to the kernel of the linear functional $f \rightarrow \hat{f}(p)$ on $\mathfrak{B}$ also belongs to the kernel of the linear functional $f \rightarrow \int f\, d\mu$ on $\mathfrak{B}$ (Theorem 3.4). Since $\mu \neq 0$, it follows there is a real number α such that $\hat{f}(p) = \alpha \int f\, d\mu$ for all $f \in \mathfrak{B}$. Now by a result of Lorch [7, Theorem 15] $p \in X$. Thus in the present circumstances every isolated point of Y belongs to X. As already pointed out in 7.1, each point of X is an isolated point of Y. Combining these remarks, X is precisely the set of isolated points of Y.

8.3. This last statement is reminiscent of a remark of Kaplan [6, p. 75]. This remark is to the effect that an *arbitrary* bounded function on X gives rise to an element of the bidual of $\mathfrak{C}$ in a natural way. From our point of view, the corresponding statement, which we shall proceed to verify, is that X is discrete and C^*-embedded in Y, for the δ-topology. (To say X is C^*-embedded in Y is to say every bounded continuous function on X admits a continuous extension to Y.) Recalling the preceding paragraph, we see that X is δ-discrete and δ-open in Y. It is known that every open subspace of an extremally disconnected space is C^*-embedded [4, Exercise 1H6]; hence we need only show that Y is extremally disconnected. Recall that in the dual of a Banach lattice every bounded subset has a least upper bound ([1, Chap. XV, Theorem 10] and [2, Chap. II, § 2, Theorem 1]), and hence $C^*(Y)$ has this property. It follows now by [4, Exercises 3N6 and 6M1], that Y is extremally disconnected. Thus it is in fact true that X is discrete and C^*-embedded in Y. (The references are all to standard books; the history of the theorems used may be traced through these books.)

It remains to show how, given an arbitrary bounded function ϕ on X, we can obtain the precise element of the bidual obtained by Kaplan from ϕ. The procedure is as follows. Using the δ-topology throughout, we define a function Φ on $\mathrm{cl}_Y X$ by extending ϕ by continuity, and then we set $\Phi(p) = 0$ for each $p \in Y \backslash \mathrm{cl}_Y X$. Since the closure of the open set X in the extremally disconnected space Y is open, the Φ so defined is δ-continuous. We omit the verification that Φ, when regarded as an element of the bidual, is that element considered by Kaplan; the primary point we are making here is that some definite element of the bidual is obtained.

8.4. It is of some interest to contrast our δ-topology with the ι-topology defined on $\hat{X}$ by Lorch [7, Definition 4]. This latter topology may, for our purposes, most readily be described as the coarsest topology on X for which $\hat{X}$ is P-space, and which is finer than the β-topology; a P-space being a topological space in which every G_δ is open. As shown by Lorch [7, Theorem 8], the ι-topology is the coarsest topology on $\hat{X}$ for which the Baire functions of $\hat{X}$ are continuous. Note particularly that we here consider the Baire functions of $\hat{X}$, not those of X as we did before.

We first consider the relative topologies induced in some fixed X_μ by the δ- and ι-topologies. We know that on X_μ the δ- and β-topologies coincide. Hence since the ι-topology is finer than the β-topology, on X_μ the ι-topology is finer than the δ-topology. Towards determining whether it is strictly finer or not, note first that, since every subspace of a P-space is a P-space, X_μ is a P-space for the ι-topology. Then note also that X_μ is compact for the δ-topology (Corollaries 3.3 and 5.3). Since it is true that a compact P-space is always finite [4, Exercise 4K1], if the ι- and δ-topologies coincide on X_μ, X_μ is finite. It follows in particular, since X_μ is δ-open in Y, that each point of X_μ is an isolated point of Y; hence $X_\mu \subset X$. Now suppose that $x_1, \cdots, x_n$ are distinct points of X_μ. Then by Theorem 3.4, the linear functional $f \to \int f\, d\mu$ on B vanishes whenever all the functionals $f \to f(x_i)$ do. Thus the first functional is a linear combination of the latter ones; i. e. there are real-numbers $\alpha_1, \cdots, \alpha_n$ such that $\int f\,d\mu = \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n)$ for all $f \in \mathfrak{B}$. To recapitulate, unless μ is an especially simple sort of measure, on X_μ the ι-topology must be strictly finer than the δ-topology.

Now consider the relative ι- and δ-topologies on X. As we have already mentioned earlier in this section, X is discrete for the δ-topology. But the ι-topology on X need not be discrete. Thus on X the δ-topology is finer than the ι-topology and is sometimes strictly finer. To summarize, there are cases in which the δ-topology is neither finer nor coarser than the ι-topology.

STATE UNIVERSITY OF NEW YORK AT ALBANY
ALBANY, N. Y.

REFERENCES.

[1] G. Birkhoff, *Lattice theory*, revised edition, American Mathematical Society Colloquium Publications, vol. XXV, American Mathematical Society, New York, 1948.

[2] N. Bourbaki, *Intégration*, chap. I-IV, Hermann, Paris, 1952.

[3] ———, *Topologie générale*, chap. 1-2, 3rd ed., Hermann, Paris, 1961.

[4] L. Gillman and M. Jerison, *Rings of continuous functions*, Van Nostrand, Princeton, 1960.

[5] S. Kabutani, "Concrete representation of abstract (M)-spaces," *Annals of Mathematics*, vol. 42 (1941), pp. 994-1024.

[6] S. Kaplan, "On the second dual of the space of continuous functions," *Transactions of the American Mathematical Society*, vol. 86 (1957), pp. 70-90.

[7] E. R. Lorch, "Compactification, Baire functions, and Daniel integration," *Acta Scientiarum Mathematicarum*, vol. XXIV (1963), pp. 204-218.

[8] H. Nakano, *Modern spectral theory*, Maruzen Co., Tokyo, 1950.

AN ANALYTIC PROOF OF THE ROGERS-RAMANUJAN-GORDON IDENTITIES.

By George E. Andrews.

1. Introduction. In [2], Gordon gives a combinatorial proof of the following beautiful generalization of the Rogers-Ramanujan identities.

Theorem. *Let a and k be integers with $0 < a \leqq k$. Let $A_{k,a}(n)$ denote the number of partitions of n into parts not of the forms $(2k+1)m$, $(2k+1)m \pm a$; $A_{k,a}(0) = 1$. Let $B_{k,a}(n)$ denote the number of partitions of n of the form*

$$n = b_1 + \cdots + b_s$$

with $b_i \geqq b_{i+1}$; $b_i - b_{i+k-1} \geqq 2$, and with 1 appearing as a summand at most $a-1$ times; $B_{k,a}(0) = 1$. Then

$$A_{k,a}(n) = B_{k,a}(n).$$

Gordon states, however, [2; p. 394] that he has been unable to deduce this theorem from the corresponding identities of Alder [1] which generalize the Rogers-Ramanujan identities analytically.

The object of this paper is to give an analytic proof of Gordon's theorem along the lines of Ramanujan's proof of the original identities [4].

2. Proof of the theorem. We define

$$C_{k,i}(x) = 1 - x^i q^i + \sum_{\mu=1}^{\infty} (-1)^{\mu} x^{k\mu} q^{\frac{1}{2}(2k+1)\mu(\mu+1)-i\mu}(1 - x^i q^{(2\mu+1)i}) \frac{(1-xq)\cdots(1-xq^{\mu})}{(1-q)\cdots(1-q^{\mu})}.$$

Selberg [5; p. 4, equation 3] has proved

$$C_{k,i}(x) = C_{k,i-1}(x) + x^{i-1}q^{i-1}(1-xq)C_{k,k-i+1}(xq). \tag{2.1}$$

If we define

$$Q_{k,i}(x) = C_{k,i}(x) \prod_{j=1}^{\infty} (1 - xq^j)^{-1},$$

then

Received July 8, 1965.

$$Q_{k,i}(x) = Q_{k,i-1}(x) + x^{i-1}q^{i-1}Q_{k,k-i+1}(xq). \tag{2.2}$$

We may expand $Q_{k,i}(x)$ as follows

$$Q_{k,i}(x) = \sum_{N=-\infty}^{\infty} \sum_{M=-\infty}^{\infty} c_{k,i}(M,N)x^M q^N, \ |x| \leqq 1, |q| < 1. \tag{2.3}$$

We then easily verify by means of the definition of $Q_{k,i}(x)$ and (2.2) that

$$c_{k,0}(M,N) = 0 \text{ for all } k,\ M,\ N. \tag{2.4}$$

$$c_{k,i}(M,N) = \begin{cases} 1 \text{ if } M = 0 \text{ and } N = 0,\ 1 \leqq i \leqq k \\ 0 \text{ if either } M \leqq 0 \text{ or } N \leqq 0 \text{ and not both are } 0 \end{cases} \tag{2.5}$$

$$c_{k,i}(M,N) - c_{k,i-1}(M,N) = c_{k,k-i+1}(M-i+1, N-M),\ 1 \leqq i \leqq k. \tag{2.6}$$

One easily verifies by methematical induction that the $c_{k,i}(M,N)$ are uniquely determined by (2.4), (2.5), and (2.6).

Let $p_{k,i}(M,N)$ denote the number of partitions of N into M parts of the form $N = b_1 + \cdots + b_M$, with $b_i \geqq b_{i+1}$, $b_i - b_{i+k-1} \geqq 2$, and 1 appearing as a summand at most $i-1$ times. The $p_{k,i}(M,N)$ clearly satisfy (2.4) and (2.5). We now show that they satisfy (2.6).

$p_{k,i}(M,N) - p_{k,i-1}(M,N)$ enumerates the number of partitions of the form given in the previous paragraph with the added condition that 1 appears *exactly* $i-1$ times as a summand. We note that if 1 appears exactly $i-1$ times then 2 can appear at most $k-i$ times. Let us now subtract 1 from every summand of the partition under consideration. Since 1 appeared exactly $i-1$ times we have reduced the number of summands to $M-i+1$. Since we have subtracted M ones from our partition, we are now partitioning $N-M$. Since 2 could appear at most $k-i$ times formerly, now 1 appears at most $k-i$ times. Thus we now have a partition of the form enumerated by $p_{k,k-i+1}(M-i+1, N-M)$. The above procedure clearly establishes a one-to-one correspondence between the partitions enumerated by $p_{k,i}(M,N) - p_{k,i-1}(M,N)$ and the partitions enumerated by $p_{k,k-i+1}(M-i+1, N-M)$. Therefore

$$p_{k,i}(M,N) - p_{k,i-1}(M,N) = p_{k,k-i+1}(M-i+1, N-M).$$

Thus the $p_{k,i}(M,N)$ fulfill (2.4), (2.5), and (2.6). Therefore by the remark following (2.6) we have

$$p_{k,i}(M,N) = c_{k,i}(M,N). \tag{2.7}$$

Thus

$$\sum_{N=0}^{\infty} A_{k,a}(N)\, q^N = \prod_{n\not\equiv 0, \pm a \pmod{2k+1}}^{\infty} (1-q^n)^{-1}$$

$$= Q_{k,a}(1) = \sum_{N=-\infty}^{\infty} \sum_{M=-\infty}^{\infty} p_{k,a}(M,N)\, q^N = \sum_{N=0}^{\infty} B_{k,a}(N)\, q^N,$$

where the second equality follows from Jacobi's identity [3; p. 282]. Therefore $A_{k,a}(N) = B_{k,a}(N)$.

THE PENNSYLVANIA STATE UNIVERSITY,
UNIVERSITY PARK, PENNSYLVANIA.

REFERENCES.

[1] H. L. Alder, "Generalizations of the Rogers-Ramanujan identities," *Pacific Journal of Mathematics*, vol. 4 (1954), pp. 161-168.

[2] B. Gordon, "A combinatorial generalization of the Rogers-Ramanujan identities," *American Journal of Mathematics*, vol. 83 (1961), pp. 393-399.

[3] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 4th ed., Oxford University Press, Oxford, 1960.

[4] S. Ramanujan, "Proof of certain identities in combinatory analysis," *Proceedings of the Cambridge Philosophical Society*, vol. 19 (1919), pp. 214-216.

[5] A. Selberg, "Uber einige arithmetische Identitäten," *Avhandlinger Norske Akademie*, 1936, no. 8.

REPRESENTATIVE ALGEBRAIC STRUCTURES ON COMPLEX ANALYTIC HOMOGENEOUS SPACES.*

By G. HOCHSCHILD and G. D. MOSTOW.

Introduction. Let G be a complex analytic linear group, and let L be a closed complex Lie subgroup of G. In [5], we have characterized the pairs (G, L) such that there are quasi-affine algebraic variety structures on $L\backslash G$ compatible with its structure as a complex analytic manifold and with the translation action of G, and we have seen that all such structures arise in the natural way from algebraic homogeneous spaces $K\backslash G^*$, where G^* is an algebraic group hull of G and K is an algebraic subgroup of G^* such that $K \cap G = L$ and $KG = G^*$. Our purpose here is to study the more general, not necessarily quasi-affine, algebraic variety structures on homogeneous spaces $L\backslash G$ that arise from such algebraic homogeneous spaces $K\backslash G^*$. In fact, by Theorems 2.2 and 2.5, the structures thus obtained are the *representative* algebraic structures as defined in Definition 2.1, which refers only to the algebra $\boldsymbol{R}(G)$ of all complex analytic representative functions on G.

Section 1 is of a preliminary nature and begins with some general (and elementary) representation-theoretical facts which are of interest in connection with homogeneous spaces. It also contains the statements of the basic algebraic-geometric facts on algebraic homogeneous spaces that are used later on. Moreover, we have included a simple proof (for the case of algebraic *linear* groups) of the known fact that an algebraic homogeneous space is quasi-projective, whence every representative algebraic structure is quasi-projective.

It turns out that all those results of [5] and [6] which are not directly concerned with sub*algebras* of $\boldsymbol{R}(G)$ generalize to representative algebraic structures. In particular, the criterion for the existence of such structures is the expected modification of Theorem 7.1 of [5] (Theorem 3.3 here). One of our present results, Theorem 5.1, was not brought out before even for quasi-affine structures. It says that, if the G-action is ignored, all representative algebraic structures on a given $L\backslash G$ are isomorphic. Thus our present investigation has a wider scope, not in that we consider more general

---

Received August 22, 1965.

* Research done with partial support of N. S. F. Grants GP 1610 and GP 1662.

structures on the same homogeneous spaces, but in that we consider a larger class of homogeneous spaces.

Many of our proofs here are based on results of [5] and [6]. New aspects arise from dealing with fields of rational functions rather than with algebras of polynomial functions. In some connections, this change has the effect of trivializing problems met in dealing with quasi-affine structures; in other connections, the difficulties are increased.

1. Algebraic homogeneous spaces.

We shall need an elementary representation-theoretical result from [1] (p. 132), which was not singled out there, so that we state it explicitly here. The proof is so short that we reproduce it for the convenience of the reader.

LEMMA 1.1. *Let P be an algebraic linear group over an arbitrary field F, and let Q be an algebraic subgroup of P. Let N be a finite-dimensional rational Q-module, and denote the transform of an element n of N by an element q of Q by $q \cdot n$. Then there is a rational homomorphism ρ of Q into the multiplicative group of F and a finite-dimensional rational P-module M, with module structure $(p,m) \rightarrow p(m)$, such that N is contained as an F-subspace in M and $q(n) = \rho(q) q \cdot n$ for every q in Q and every n in N.*

Proof. Let S denote the space of the rational representative functions on Q that is associated with the given rational Q-module N. If we regard S as a rational Q-module for the left action of Q given by $(q,s) \rightarrow q \cdot s$, where $(q \cdot s)(q_1) = s(q_1 q)$, then N may be identified with a Q-submodule of the direct sum of a finite number of copies of S. Since every rational representative function on Q is the restriction to Q of a rational representative function on P, we see from this that N may be written in the form U/V, where U is a Q-submodule of some finite-dimensional rational P-module T and V is a Q-submodule of U.

Let n be the dimension of V, and let $E^{n+1}(T)$ denote the homogeneous component of degree $n+1$ of the exterior F-algebra built over T, regarded as a rational P-module in the canonical fashion. This contains the exterior product $UE^n(V)$ as a Q-submodule. Evidently, this Q-module $UE^n(V)$ is isomorphic with the tensor product Q-module $N \otimes_F E^n(V)$. Now $E^n(V)$ is 1-dimensional, and the action of an element q of Q on $E^n(V)$ is simply the scalar multiplication by the determinant of the linear automorphism of V that corresponds to q. Hence it is clear that if u is any non-zero element of $E^n(V)$ then the injection $n \rightarrow n \otimes u$, followed by the Q-module injection of $N \otimes_F E^n(V)$ into $E^{n+1}(T)$, satisfies the requirements of Lemma 1.1.

Observe that this proof shows that *Lemma* 1.1 *actually holds for the representation theory of any category of groups, provided that one adds the assumption* (which is always satisfied in the algebraic case) *that the representative functions associated with N be extendible to representative functions on P.*

With P and Q as above, we denote by $\boldsymbol{R}(P)$ the F-algebra of all rational representative functions on P. If P is irreducible then $\boldsymbol{R}(P)$ is an integral domain, and we can form the field of fractions of $\boldsymbol{R}(P)$, which we denote by $[\boldsymbol{R}(P)]$. The left P-module structure of $\boldsymbol{R}(P)$, in which P acts by algebra automorphisms on $\boldsymbol{R}(P)$, extends canonically to a left P-module structure on $[\boldsymbol{R}(P)]$ in which P acts by field automorphisms on $[\boldsymbol{R}(P)]$. We denote the Q-fixed subfield of $[\boldsymbol{R}(P)]$ by $[\boldsymbol{R}(P)]^Q$.

PROPOSITION 1.2. *Let P be an irreducible algebraic linear group over an arbitrary field F, and let Q be an algebraic subgroup of P. Let f be any element of $[\boldsymbol{R}(P)]^Q$. Then f is the ratio of two Q-semi-invariants in $\boldsymbol{R}(P)$, i. e., there is a rational homomorphism ρ of Q into the multiplicative group of F, and there are elements s and t in $\boldsymbol{R}(P)$ such that $f=s/t$ and, for every q in Q, $q\cdot s=\rho(q)s$ and $q\cdot t=\rho(q)t$.*

Proof. Consider the rational Q-module $\boldsymbol{R}(P)f\cap\boldsymbol{R}(P)$. If $f\neq 0$ this contains a non-zero simple Q-submodule V. Let V^* denote the Q-module dual to V. We may apply Lemma 1.1 to the rational Q-module V^*. Let ρ and M be as described in Lemma 1.1. Clearly, V^* is a simple Q-module also with respect to the module structure of M. Since M is isomorphic with a P-submodule of a direct sum of a finite number of copies of $\boldsymbol{R}(P)$, it follows that there is a Q-module monomorphism η of V^*, regarded as a Q-submodule of M, into $\boldsymbol{R}(P)$. With the original Q-module structure of V^*, we have $q\cdot\eta(\alpha)=\rho(q)\eta(q\cdot\alpha)$ for every q in Q and every α in V^*.

Now let $(v_1,\cdots,v_n)$ be an F-basis of V such that $v_1(1)\neq 0$ while $v_i(1)=0$ for every other index i, where 1 stands for the neutral element of P. Let $(\alpha_1,\cdots,\alpha_n)$ be the dual basis of V^*, so that $\alpha_i(v_j)=\delta_{ij}$. Put $g_i=\eta(\alpha_i)$, so that each g_i is a non-zero element of $\boldsymbol{R}(P)$. Choose an element p in P such that $g_1(p)\neq 0$, and put $s=\sum_{i=1}^{n}(g_i\cdot p)v_i$, where $g_i\cdot p$ is the element of $\boldsymbol{R}(P)$ given by $(g_i\cdot p)(x)=g_i(px)$ for every x in P. For q in Q, write $q\cdot v_i=\sum_{j=1}^{n}h_{ji}(q)v_j$. Then we have $q\cdot g_i=\rho(q)\eta(q\cdot\alpha_i)$ $=\rho(q)\sum_{j=1}^{n}h_{ij}(q^{-1})g_j$. Hence we obtain

$$q\cdot s=\sum_{i=1}^{n}(q\cdot g_i\cdot p)\,q\cdot v_i=\sum_{ijk}\rho(q)h_{ij}(q^{-1})(g_j\cdot p)h_{ki}(q)v_k=\rho(q)s.$$

Now s lies in $\boldsymbol{R}(P)f\cap\boldsymbol{R}(P)$, and $s(1)=g_1(p)v_1(1)\neq 0$. Thus s is a non-zero Q-semi-invariant in $\boldsymbol{R}(P)$, and $s=tf$ with t in $\boldsymbol{R}(P)$. Since f is Q-fixed, we have also $q\cdot t=\rho(q)t$ for every q in Q. Since $s\neq 0$, we have $t\neq 0$ and $f=s/t$, which proves Proposition 1.2.

It is clear from the above proof and from what we observed concerning the more general applicability of Lemma 1.1 that Proposition 1.2 *holds also if P is a real or complex analytic group, Q a real or complex closed Lie subgroup of P, and $\boldsymbol{R}(P)$ stands for the algebra of the real or complex analytic representative functions on P. The homomorphism ρ is then real or complex analytic.*

Now let P be an irreducible algebraic linear group over an *algebraically closed* field F, and let Q be an algebraic subgroup of P. Let $Q\backslash P$ denote the set of cosets Qp, with p in P. Let ${}^{Q}[\boldsymbol{R}(P)]$ denote the subfield of $[\boldsymbol{R}(P)]$ consisting of the elements that are fixed under the action of Q *from the right* (derived from the action $f\rightarrow f\cdot q$ on $\boldsymbol{R}(P)$, where $(f\cdot q)(p)=f(qp)$). We recall that there is a unique algebraic variety structure on $Q\backslash P$ satisfying the following conditions: (1) the canonical map $\pi\colon P\rightarrow Q\backslash P$ is a morphism of the algebraic variety P onto the algebraic variety $Q\backslash P$; (2) if ρ is any morphism of the algebraic variety P into an algebraic variety V such that ρ is constant on each coset Qp then there is one and only one morphism ρ^{Q} of the algebraic variety $Q\backslash P$ into V such that $\rho^{Q}\circ\pi=\rho$. Moreover, the cohomomorphism π^* of π maps the field of the rational functions on $Q\backslash P$ isomorphically onto ${}^{Q}[\boldsymbol{R}(P)]$, and a rational function f on $Q\backslash P$ is defined at a point Qp of $Q\backslash P$ whenever $\pi^*(f)$ is defined at p. Thus f is defined at a point x of $Q\backslash P$ whenever there is a point p in P and elements u and v of $\boldsymbol{R}(P)$ such that $\pi(p)=x$, $v(p)\neq 0$ and $\pi^*(f)=u/v$. Finally, ${}^{Q}[\boldsymbol{R}(G)]$ separates the points of $Q\backslash P$, i.e., for every pair (x,y) of distinct points of $Q\backslash P$, there is an element g of ${}^{Q}[\boldsymbol{R}(G)]$ such that g is defined at x and at y and $g(x)\neq g(y)$. Throughout, the algebraic-geometric language we have in mind is that of [2], and we refer the reader to [3] for proofs of the above assertions.

An algebraic variety is called *quasi-projective* if it is isomorphic with an open subvariety of a projective variety. It is known that *the above algebraic variety $Q\backslash P$ is quasi-projective.* In fact, this result is a special case of a general theorem, due to W. L. Chow, on projective embeddings of algebraic varieties that are homogeneous with respect to an arbitrary algebraic group, [4]. Since the proof of the general theorem requires rather heavy machinery of algebraic geometry, we exhibit the following simple proof of our special case, which utilizes Proposition 1.2.

Let $(f_1, \cdots, f_n)$ be a set of field generators for ${}^{Q}[\boldsymbol{R}(P)]$ over F. By Proposition 1.2 (with left and right interchanged), each f_i can be written as the ratio of two right Q-semi-invariants in $\boldsymbol{R}(P)$. An evident multiplicative adjustment of numerators and denominators yields right Q-semi-invariants $u_0, \cdots, u_n$ in $\boldsymbol{R}(P)$, *all of the same 'weight'* ρ, such that $f_i = u_i/u_0$ for each i. Let S be the right P-submodule of $\boldsymbol{R}(P)$ that is generated by the u_i's, and let T be the direct sum of $n+1$ copies of S. Let t be the point $(u_0, \cdots, u_n)$ of T. Let T^* denote the projective space consisting of the one-dimensional subspaces of T, and let t^* be the point of T^* that corresponds to t. The right action of P by linear automorphisms of T induces a right action of P by algebraic variety automorphisms on T^*. Evidently, t^* is kept fixed by every element of Q. Conversely, if p is an element of P leaving t^* fixed then $(u_i/u_0) \cdot p = u_i/u_0$ for each i, whence $f \cdot p = f$ for every f in ${}^{Q}[\boldsymbol{R}(P)]$. Since ${}^{Q}[\boldsymbol{R}(P)]$ separates the points of $Q \backslash P$, this implies that p belongs to Q. Thus Q is precisely the fixer of t^* in P.

Now the map $\alpha \colon P \rightarrow T^*$ such that $\alpha(p) = t^* \cdot p$ is evidently a morphism of algebraic varieties. Since it is constant on each coset Qp, it induces a morphism $\alpha^Q \colon Q \backslash P \rightarrow T^*$ such that $\alpha^Q \circ \pi = \alpha$. Evidently, α^Q is a P-morphism with respect to the canonical actions of P on $Q \backslash P$ and on T^*. Since Q is precisely the fixer of t^* in P, it is clear that α^Q is injective. The closure, A say, of the P-orbit of t^* in T^* is a projective algebraic variety, and the image of α^Q contains an open subset of A. Because of the P-homogeneity, it follows that the image of α^Q is actually an open subvariety B of A, and thus a quasi-projective variety. Thus α^Q may be regarded as a bijective morphism β of $Q \backslash P$ onto the quasi-projective algebraic variety B. By the choice of the u_i's, the cohomomorphism of β maps the field of the rational functions on B isomorphically onto the field ${}^{Q}[\boldsymbol{R}(P)]$ of the rational functions on $Q \backslash P$. Since β is a P-morphism, it follows from the P-homogeneity that β is actually an isomorphism of algebraic varieties, so that our proof is now complete.

2. Representative algebraic structures. Let G be a complex analytic group, and let $\boldsymbol{R}(G)$ denote the algebra, over the field C of the complex numbers, of all complex analytic representative functions on G. If S is any integral domain we shall denote by $[S]$ the field of fractions of S.

Let f be an element of $[\boldsymbol{R}(G)]$. The *domain*, D_f, of f in G is the set of all elements x of G such that f may be written as a fraction a/b with a and b in $\boldsymbol{R}(G)$ and $b(x) \neq 0$. Then f may be regarded as a complex-valued function on D_f, its value at a point x of D_f being equal to $a(x)/b(x)$ for *every* pair (a, b) of elements of $\boldsymbol{R}(G)$ such that $f = a/b$ and $b(x) \neq 0$.

Now let L be a closed complex Lie subgroup of G, and let us consider the subfield ${}^{L}[\boldsymbol{R}(G)]$ of the right L-fixed elements of $[\boldsymbol{R}(G)]$. If f is an element of ${}^{L}[\boldsymbol{R}(G)]$ then then $LD_f = D_f$ and f is constant on the coset Lx for every x in D_f. Thus f determines a complex-valued function on the canonical image of D_f in $L \backslash G$, which we shall call the *domain of f in $L \backslash G$.*

We claim that *the domain in $L \backslash G$ of every element f of ${}^{L}[\boldsymbol{R}(G)]$ is a dense open subset of $L \backslash G$.* In fact, if $f = a/b$ with a and b in $\boldsymbol{R}(G)$ then D_f contains the set of non-zeros of b, which is dense in G, because b is a non-zero holomorphic function on G. Thus D_f is dense in G, and since the canonical map $G \rightarrow L \backslash G$ is surjective and continuous it follows that the domain of f in $L \backslash G$ is dense in $L \backslash G$. Clearly, D_f is open in G. Since the canonical map $G \rightarrow L \backslash G$ is an open map, it follows that the domain of f in $L \backslash G$ is open.

Definition 2.1. *Let G be a complex analytic group, L a closed complex Lie subgroup of G. A representative algebraic structure on $L \backslash G$ is the structure of an irreducible algebraic variety on $L \backslash G$ satisfying the following conditions.*

(1) *For each element x of G, the translation effected by x on $L \backslash G$ is an automorphism of the algebraic variety structure.*

(2) *Every rational function f of the algebraic variety structure is the function obtained in the canonical fashion from an element f^0 of ${}^{L}[\boldsymbol{R}(G)]$.*

Clearly, f^0 is uniquely determined by f, and the map $f \rightarrow f^0$ is a field isomorphism of the field of the rational functions onto a finitely generated left G-stable subfield of ${}^{L}[\boldsymbol{R}(G)]$, which we shall simply call *the field of the given representative algebraic structure.*

THEOREM 2.2. *Let G be a complex analytic group whose elements are separated by $\boldsymbol{R}(G)$. Let G^* be an algebraic group hull of G, and let K be an algebraic subgroup of G^* such that $KG = G^*$. Then the canonical bijection $(K \cap G) \backslash G \rightarrow K \backslash G^*$ transports the natural algebraic variety structure of $K \backslash G^*$ onto a representative algebraic structure of $(K \cap G) \backslash G$.*

Proof. As usual, denote by A the group of all proper automorphisms of $\boldsymbol{R}(G)$, (i.e., the automorphisms commuting with the right G-action).

There is a *fully stable* (i.e., stable under the translation actions of G and under the involution corresponding to the inversion map on G) finitely generated subalgebra S of $\boldsymbol{R}(G)$ such that G^* may be identified with the restriction image A_S of A, and S becomes identified with the algebra of all polynomial functions on G^*; we recall that the elements f of $\boldsymbol{R}(G)$ are viewed as complex-valued functions on A, the value of f at α being $\alpha(f)(1)$. With this identification, the rational functions on $K\backslash G^*$ are the functions obtained in the natural way from the right K-fixed elements of $[S]$, as explained in the beginning of this section, with S taking the place of $\boldsymbol{R}(G)$.

The corresponding rational functions for the transported algebraic variety structure of $(K \cap G)\backslash G$ are the functions obtained in the same way from these same elements of $[S]$, viewed as functions on G. In order to conclude that our structure on $(K \cap G)\backslash G$ satisfies condition (2) of Definition 2.1, we must show that the domains of these rational functions are the same as the domains one would obtain by viewing the corresponding elements of $[S]$ as elements of $[\boldsymbol{R}(G)]$. A priori, it is conceivable that there is an element f of $[S]$ and an element x of G such that every denominator of f in S vanishes at x, while f has a denominator in $\boldsymbol{R}(G)$ that does not vanish at x. We must show that, actually, this cannot happen.

Let T be any finitely generated fully stable subalgebra of $\boldsymbol{R}(G)$ that contains S. Then the restriction map $A_T \to A_S$ is a rational epimorphism of the algebraic linear group A_T onto the algebraic linear group A_S. Since our base field C is algebraically closed and of characteristic 0, it follows that A_S is isomorphic, as an algebraic linear group, with the factor group of A_T mod. the kernel of our epimorphism. In particular, this implies that a rational function of the algebraic variety structure of A_S is defined at a point x of A_S whenever the corresponding rational function of the algebraic variety structure of A_T is defined at a point of A_T whose restriction image in A_S is x. In other words, if an element of $[S]$ is defined at a point y of A_T when viewed as an element of $[T]$ then it is defined as it stands at the image of y in A_S. Clearly, this yields the desired result concerning the domains of the elements of $[S]$ as functions on G, when we consider the natural injection of G into A.

Thus the transported algebraic variety structure of $(K \cap G)\backslash G$ satisfies condition (2) of Definition 2.1. Since the canonical bijection $(K \cap G)\backslash G \to K\backslash G^*$ is compatible with the actions of G on $(K \cap G)\backslash G$ and on $K\backslash G^*$, it is clear that condition (1) of Definition 2.1 is also satisfied. This completes the proof of Theorem 2.2.

We wish to prove that every representative algebraic structure is of the type of Theorem 2.2. For this purpose, we require the following result.

PROPOSITION 2.3. *The rational functions of a representative algebraic structure on* $L\backslash G$ *separate the points of* $L\backslash G$.

Proof. Let u and v be distinct points of $L\backslash G$. There is an open affine subvariety M of $L\backslash G$ such that u lies in M. We have seen at the beginning of this section that the domain of every rational function of the algebraic variety structure on $L\backslash G$ is topologically an open dense subset of $L\backslash G$. Hence M is topologically open and dense in $L\backslash G$. Let H be the set of all elements x of G such that $u\cdot x$ lies in M. Since M is topologically open in $L\backslash G$, it is clear that H is open in G, whence $v\cdot H$ is a topologically open subset of $L\backslash G$. Since M is topologically dense in $L\backslash G$, it follows that $v\cdot H$ meets M. Thus there is an element x in G such that both $u\cdot x$ and $v\cdot x$ lie in M. Since M is an open affine algebraic subvariety of $L\backslash G$, it follows that there is a rational function f of the algebraic variety structure on $L\backslash G$ such that f is defined at $u\cdot x$ and at $v\cdot x$ and $f(u\cdot x)\neq f(v\cdot x)$. Now the translate $x\cdot f$ of f is a rational function of the algebraic variety structure on $L\backslash G$ (because of condition (1) of Definition 2.1), it is defined at u and at v, and its values there are distinct. Thus Proposition 2.3 is proved.

We denote by $f\rightarrow f'$ the canonical involution of $\boldsymbol{R}(G)$; $f'(x)=f(x^{-1})$. This extends uniquely to a field involution of $[\boldsymbol{R}(G)]$, which we write in the same way. If P is any subset of $[\boldsymbol{R}(G)]$ then P' denotes the image of P under this involution. If G is a complex analytic group whose points are separated by $\boldsymbol{R}(G)$ then its points are separated already by a finitely generated fully stable subalgebra S of $\boldsymbol{R}(G)$. If F is any finitely generated subfield of $[\boldsymbol{R}(G)]$ we may take S such that, moreover, F and F' lie in $[S]$. We may view A_S also as a group of field automorphisms of $[S]$, and we denote the fixer of F' in A_S by $(A_S)^{F'}$. Clearly, $(A_S)^{F'}=(A^{F'})_S$, and $(A_S)^{F'}$ is an algebraic subgroup of A_S.

PROPOSITION 2.4. *Let* G *be a complex analytic group whose points are separated by* $\boldsymbol{R}(G)$. *Let* F *be a finitely generated left* G*-stable subfield of* $[\boldsymbol{R}(G)]$, *and let* S *be a finitely generated fully stable subalgebra of* $\boldsymbol{R}(G)$ *such that* S *separates the points of* G *and* $[S]$ *contains* F *and* F'. *Then the elements of* F, *when viewed as rational functions on* A_S, *are right* $(A_S)^{F'}$-*fixed and separate the points of* $(A_S)^{F'}\backslash A_S$. *In this way,* F *may be identified with the field of the rational functions on* $(A_S)^{F'}\backslash A_S$. *Moreover,* F' *coincides with the* $A^{F'}$-*fixed subfield* $[\boldsymbol{R}(G)]^{A^{F'}}$ *of* $[\boldsymbol{R}(G)]$.

Proof. Let x and y be elements of G, and let u be an element of S. We have

$$(y^{-1}\cdot u')(x^{-1}) = u'(x^{-1}y^{-1}) = u(yx) = (x\cdot u)(y).$$

Since the canonical image G_S of G in A_S is algebraically dense in A_S, it follows that $\beta^{-1}(u')(x^{-1}) = (x\cdot u)(\beta)$ for every β in A_S, and hence that $\beta^{-1}(u')(\alpha^{-1}) = \alpha(u)(\beta)$ for all α and β in A_S, i. e., that $\beta^{-1}(u')(\alpha^{-1}) = u(\beta\alpha)$. This means that, when $\beta^{-1}(u')$ and u are regarded as functions on the algebraic linear group A_S, we have $(\beta^{-1}(u'))' = u\cdot\beta$. The same is therefore true if u is an arbitrary element of $[S]$, regarded as a rational function on A_S. If u belongs to F and β belongs to $(A_S)^{F'}$ this gives $u = u\cdot\beta$, which shows that the elements of F, when viewed as rational functions on A_S, are right $(A_S)^{F'}$-fixed, and thus may be regarded as rational functions on the algebraic variety $(A_S)^{F'}\backslash A_S$. We shall now show that, as such, they separate the points of $(A_S)^{F'}\backslash A_S$.

Suppose that α and β are elements of A_S that are not separated by F, and that f is an element of F that is defined at α and at β. This means that we have $f = a_1/b_1 = a_2/b_2$, where a_1, a_2, b_1, b_2 belong to S, $b_1(\alpha) \neq 0$ and $b_2(\beta) \neq 0$. It follows that for all elements x in some open neighborhood of 1 in G, we still have $(x\cdot b_1)(\alpha) \neq 0$ and $(x\cdot b_2)(\beta) \neq 0$. Now $x\cdot f$ lies in F, and $x\cdot f = (x\cdot a_1)/(x\cdot b_1) = (x\cdot a_2)/(x\cdot b_2)$, so that $x\cdot f$ is defined at α and at β. Since α and β are not separated by F, we have therefore

$$(x\cdot a_1)(\alpha)(x\cdot b_2)(\beta) = (x\cdot a_2)(\beta)(x\cdot b_1)(\alpha).$$

By a formula obtained at the beginning of this proof, this relation may be written

$$\alpha^{-1}(a_1')(x^{-1})\beta^{-1}(b_2')(x^{-1}) = \beta^{-1}(a_2')(x^{-1})\alpha^{-1}(b_1')(x^{-1}).$$

Since this holds for all elements x of some non-empty open subset of G, it follows that $\alpha^{-1}(a_1')\beta^{-1}(b_2') = \beta^{-1}(a_2')\alpha^{-1}(b_1')$, whence $\alpha^{-1}(f') = \beta^{-1}(f')$. Operating from the right with an arbitrary element z of G, we obtain from this that $\alpha^{-1}((z^{-1}\cdot f)') = \beta^{-1}((z^{-1}\cdot f)')$ whenever f is an element of F that is defined at α and at β. This means that $\alpha^{-1}(g') = \beta^{-1}(g')$ whenever g is an element of F for which there is an element z of G such that $z\cdot g$ is defined at α and at β, i. e., such that g is defined at αz_S and at βz_S.

Now let g be any element of F, and write $g = a/b$ with a and b in S. The translates $b\cdot x$ and $b\cdot\beta$ are polynomial functions on A_S, and the intersection of their sets of non-zeros is a Zariski-open subset of A_S. Since G_S is algebraically dense in A_S, it follows that there is an element z in G such that neither $b\cdot\alpha$ nor $b\cdot\beta$ vanishes at z_S. But this means that g is defined at

αz_S and at βz_S so that, by the above, $\alpha^{-1}(g') = \beta^{-1}(g')$. Thus we have shown that α^{-1} and β^{-1} coincide on F', which means that $(A_S)^{F'}\alpha = (A_S)^{F'}\beta$. Thus F separates the points of $(A_S)^{F'}\backslash A_S$. Since our base field C is of characteristic 0, this implies that F must coincide with the field of all rational functions on $(A_S)^{F'}\backslash A_S$.

When $[S]$ is regarded as the field of the rational functions on A_S, the field of the rational functions on $(A_S)^{F'}\backslash A_S$ is identifiable with the right $(A_S)^{F'}$-fixed subfield ${}^{(A_S)^{F'}}[S]$ of $[S]$. In this sense, we have $F = {}^{(A_S)^{F'}}[S]$. On applying the involution $f \to f'$ to this result, we obtain $F' = [S]^{(A_S)^{F'}}$, i.e., F' coincides with the $(A_S)^{F'}$-fixed subfield of $[S]$, with respect to the natural action of A_S on $[S]$. Since S may be chosen so that $[S]$ contains any given element of $[\mathbf{R}(G)]$, it follows that $F' = [\mathbf{R}(G)]^{A^{F'}}$, and our proof of Proposition 2.4 is complete.

THEOREM 2.5. *Let G be a complex analytic group whose points are separated by $\mathbf{R}(G)$, and let L be a closed complex Lie subgroup of G. Suppose that there is given a representative algebraic structure on $L\backslash G$, with field $F \subset {}^{L}[\mathbf{R}(G)]$. Let S be a finitely generated fully stable subalgebra of $\mathbf{R}(G)$ such that S separates the points of G and $[S]$ contains F and F'. Then $(A_S)^{F'}G_S = A_S$, $(A_S)^{F'} \cap G_S = L_S$, and the canonical map $L\backslash G \to (A_S)^{F'}\backslash A_S$ is an isomorphism of algebraic varieties. Moreover, we have $A^{F'}G = A$ and $A^{F'} \cap G = L$.*

Proof. Since F is right L-fixed, F' is left L-fixed, so that $L_S \subset (A_S)^{F'}$. Now suppose that x is an element of G such that x_S lies in $(A_S)^{F'}$. Then $f \cdot x = f$ for every element f of F, so that the points Lx and L of $L\backslash G$ are not separated by F. By Proposition 2.3, we have therefore $Lx = L$, i.e., $x \in L$. Thus we have shown that $L_S = (A_S)^{F'} \cap G_S$. Since the map $G \to G_S$ is injective, it follows that $A^{F'} \cap G = L$.

Since $L_S = (A_S)^{F'} \cap G_S$, the canonical map $L\backslash G \to (A_S)^{F'}\backslash A_S$ is injective. Proposition 2.4 shows that this map is a morphism of algebraic varieties whose cohomomorphism maps the field of the rational functions on $(A_S)^{F'}\backslash A_S$ isomorphically onto the field of the rational functions on $L\backslash G$. In particular, $L\backslash G$ and $(A_S)^{F'}\backslash A_S$ are therefore of the same dimension as algebraic varieties. Since the rational functions of each algebraic variety structure are locally holomorphic where they are defined, it follows that $L\backslash G$ and $(A_S)^{F'}\backslash A_S$ are of the same dimension also as complex analytic manifolds. Hence it follows from the elementary theory of complex analytic homogeneous spaces, by considering the differential of the injective complex analytic map $L\backslash G \to (A_S)^{F'}\backslash A_S$, that the image of $L\backslash G$ contains a neighborhood of the point $(A_S)^{F'}$ in

$(A_S)^{F'}\backslash A_S$. On the other hand, this image is evidently the canonical image of the subgroup $(A_S)^{F'}G_S$ of A_S. Hence we conclude that $(A_S)^{F'}G_S$ is an open subgroup of A_S, and thus coincides with A_S. Since the kernel of the restriction epimorphism $A \to A_S$ is contained in $A^{F'}$, it follows that $A^{F'}G = A$.

Now we know that the map $L\backslash G \to (A_S)^{F'}\backslash A_S$ is a *bijective* morphism of algebraic varieties and that its cohomomorphism is an isomorphism of the field of rational functions of $(A_S)^{F'}\backslash A_S$ onto the field of rational functions of $L\backslash G$. Since our map is compatible with the natural actions of G on $L\backslash G$ and on $(A_S)^{F'}\backslash A_S$, a familiar homogeneity argument shows that the inverse of our bijective morphism is also a morphism of algebraic varieties, so that Theorem 2.5 is proved.

THEOREM 2.6. *Let G be a complex analytic group whose points are separated by $\mathbf{R}(G)$, and let F be a finitely generated left G-stable subfield of $[\mathbf{R}(G)]$ such that $A^{F'}G = A$. Then F is the field of a representative algebraic structure on $(A^{F'} \cap G)\backslash G$.*

Proof. Let S be as in Proposition 2.4. Since $A^{F'}G = A$, we have $(A^{F'})_S G_S = A_S$, i.e., $(A_S)^{F'}G_S = A_S$. We know from Proposition 2.4 that F may be identified in the natural way with the field of all rational functions on $(A_S)^{F'}\backslash A_S$. The injection $G_S \to A_S$ induces an injection of $((A_S)^{F'} \cap G_S)\backslash G_S$ into $(A_S)^{F'}\backslash A_S$, which is surjective because $(A_S)^{F'}G_S = A_S$. Since the kernel of the restriction epimorphism $A \to A_S$ lies in $A^{F'}$, it is clear that $(A_S)^{F'} \cap G_S = (A^{F'} \cap G)_S$. Since S separates the points of G, we may identify G_S with G and $(A^{F'} \cap G)_S$ with $A^{F'} \cap G$. Thus the canonical map $(A^{F'} \cap G)\backslash G \to (A_S)^{F'}\backslash A_S$ is bijective and transports the algebraic variety structure of $(A_S)^{F'}\backslash A_S$ onto an algebraic variety structure on $(A^{F'} \cap G)\backslash G$ whose field of rational functions is F. By Theorem 2.2, this is a representative algebraic structure, so that Theorem 2.6 is proved.

3. Existence criterion.

We shall need the following easy generalization of Theorem 6.1 of [5].

THEOREM 3.1. *Let G be a complex analytic group, L a closed complex Lie subgroup of G. Suppose that ${}^L[\mathbf{R}(G)]$ separates the points of $L\backslash G$. Then there is a finite subset of ${}^L[\mathbf{R}(G)]$ that separates the points of $L\backslash G$.*

Proof. We may evidently assume, without loss of generality, that $\mathbf{R}(G)$ separates the points of G. Let G' denote the commutator subgroup of G. Then G' has the structure of an algebraic linear group such that the restriction to G' of every complex analytic representation of G is a rational representation of G' (see the proof of Theorem 6.1 in [5]).

With slight changes in the auxilary definitions, which are dictated by our dealing with subsets of $[\boldsymbol{R}(G)]$ rather than with subsets of $\boldsymbol{R}(G)$, the relevant part of the proof of Theorem 6.1 in [5] applies here to show that that $L \cap G'$ is an algebraic subgroup of G', that LG' is closed in G, and that a finite subset X of ${}^{LG'}\boldsymbol{R}(G)$ separates the points of $(LG')\backslash G$. If we restrict the elements of ${}^{L}[\boldsymbol{R}(G)]$ to G' we obtain a field of right $(L \cap G')$-fixed rational functions on the algebraic linear group G' that separates the points of $(L \cap G')\backslash G'$. Hence this field, when regarded as a field of functions on $(L \cap G')\backslash G'$, becomes identified with the field of *all* rational functions on the algebraic variety $(L \cap G')\backslash G'$. Now we observe, generally, that if V is an algebraic variety whose points are separated by the field F of the rational functions on V then the points of V are separated by a finite subset of F. In order to see this, consider the product variety $V \times V$. For each finite subset T of F, let $M(T)$ denote the set of all points (x, y) of $V \times V$ such that x and y are separated by an element of T. Evidently, each $M(T)$ is a Zariski-open subset of $V \times V$. Since an algebraic variety satisfies the maximal condition for open subsets, there is a finite subset T_0 of F such that $M(T_0)$ is maximal in the family of all $M(T)$'s, and since F separates the points of V it is clear that $M(T_0)$ must be the complete complement of the diagonal in $V \times V$, so that T_0 separates the points of V.

Returning to $(L \cap G')\backslash G'$, we may now conclude that there is a finite subset Y of ${}^{L}[\boldsymbol{R}(G)]$ that separates the points of $(L \cap G')\backslash G'$, and thus the points of $L\backslash(LG')$. Now it is clear that the finite subset $X \cup Y$ of ${}^{L}[\boldsymbol{R}(G)]$ separates the points of $L\backslash G$, which completes the proof of Theorem 3.1.

LEMMA 3.2. *Let G be a complex analytic group whose points are separated by $\boldsymbol{R}(G)$. Let X and Y be closed complex Lie subgroups of G such that $X\backslash G$ and $Y\backslash G$ have representative algebraic structures and $XY = G$. Then $(X \cap Y)\backslash G$ has a representative algebraic structure.*

Proof. By Theorem 2.5, there is a finitely generated fully stable subalgebra S of $\boldsymbol{R}(G)$ satisfying the following conditions: S separates the points of G; the group A_S contains algebraic subgroups X^* and Y^* such that $X^* \cap G_S = X_S$, $Y^* \cap G_S = Y_S$, $X^*G_S = A_S$, $Y^*G_S = A_S$. We have $G_S = X_S Y_S$, so that, from $X^*G_S = A_S$, we get $X^*Y_S = A_S$. A fortiori, $X^*Y^* = A_S$.

Now consider the canonical map $(X \cap Y)\backslash G \rightarrow (X\backslash G) \times (Y\backslash G)$. This map is evidently injective. From the fact that $XY = G$, one sees readily that this map is also surjective. Similarly, we see that the canonical map $(X^* \cap Y^*)\backslash A_S \rightarrow (X^*\backslash A_S) \times (Y^*\backslash A_S)$ is bijective. Now consider the following commutative diagram of canonical maps:

$$\begin{tikzcd}(X \cap Y)\backslash G \arrow[r] \arrow[d] & (X\backslash G) \times (Y\backslash G) \arrow[d] \\ (X^* \cap Y^*)\backslash A_S \arrow[r] & (X^*\backslash A_S) \times (Y^*\backslash A_S).\end{tikzcd}$$

We have just seen that the horizontal maps are bijective. The vertical map on the right is defined from the bijective canonical maps $X\backslash G \rightarrow X^*\backslash A_S$ and $Y\backslash G \rightarrow Y^*\backslash A_S$, and is therefore bijective. It follows that the vertical map on the left is bijective. By Theorem 2.2, we therefore obtain a representative algebraic structure on $(X \cap Y)\backslash G$ by transporting the natural algebraic variety structure of $(X^* \cap Y^*)\backslash A_S$. This completes the proof of Lemma 3.2.

Let α denote the adjoint representation of G, and denote the algebraic group hull of $\alpha(G)$ in the full linear group on the Lie algebra $\mathbf{L}(G)$ of G by $\alpha(G)^*$. We note, as at the beginning of Section 7 in [5], that $\alpha(G)^*$ lies in the canonical image of the group $\mathbf{A}(G)$ of all complex analytic automorphisms of G in the full linear group on $\mathbf{L}(G)$. We denote by $\mathbf{S}(L)$ the subgroup of $\alpha(G)^*$ consisting of those elements of $\alpha(G)^*$ which correspond to elements of $\mathbf{A}(G)$ leaving the subgroup L of G stable. The normalizer of L in G is denoted by $N_G(L)$.

THEOREM 3.3. *Let L be a closed complex Lie subgroup of the complex analytic linear group G. Then $L\backslash G$ has a representative algebraic structure if and only if the following three conditions are satisfied:*

(1) ${}^L[\mathbf{R}(G)]$ *separates the points of* $L\backslash G$

(2) $\mathbf{S}(L)\alpha(G) = \alpha(G)^*$;

(3) $N_G(L)/L$ *has only a finite number of connected components.*

Proof. After an appeal to Theorem 2.5, the proof of the necessity of the above conditions is identical with the corresponding part of the proof of Theorem 7.1 in [5].

In order to prove the sufficiency of these conditions, we show first that L may be assumed to have only a finite number of connected components. By condition (1) and Theorem 3.1, we can find a finitely generated fully stable subalgebra S of $\mathbf{R}(G)$ such that S separates the points of G and ${}^L[S]$ separates the points of $L\backslash G$. Let L^* denote the algebraic group hull of L_S in A_S. Exactly as in the corresponding part of the proof of Theorem 7.1 in [5], one sees that L^*G_S is connected, and thus a complex analytic subgroup of A_S, that the pair (L^*G_S, L^*) satisfies the three conditions of our theorem, and that the canonical map $L\backslash G \rightarrow L^*\backslash(L^*G_S)$ is bijective. Since L^* is an algebraic linear group, it has only a finite number of connected components. Hence, if our theorem holds for subgroups with finite component groups, we have a representative algebraic structure on $L^*\backslash(L^*G_S)$. The above bijection transports this structure onto an algebraic variety structure on $L\backslash G$, which evidently satisfies condition (1) of Definition 2.1. The rational functions of this structure are the restrictions to G of the rational functions for $L^*\backslash(L^*G_S)$. The domain of such a function in G is the domain obtained in the canonical fashion by regarding it as an element of $[T]$, where T is the restriction image of $\boldsymbol{R}(L^*G_S)$ in $\boldsymbol{R}(G)$. Clearly, T is a fully stable subalgebra of $\boldsymbol{R}(G)$ and separates the points of G. From our proof of Theorem 2.2, we know that the domain of an element of $[T]$ coincides with its domain as an element of $[\boldsymbol{R}(G)]$. Thus we conclude that our algebraic variety structure on $L\backslash G$ satisfies also the condition (2) of Definition 2.1, so that it is a representative algebraic structure. Thus it suffices to prove Theorem 3.3 in the case where L has only a finite number of connected components.

In that case, the main part of the proof of Theorem 7.1 in [5] applies when one begins with a finitely generated fully stable subalgebra S of $\boldsymbol{R}(G)$ such that S separates the points of G and contains $\mathrm{Hom}(G, C)$, and the fixer in G of $[S]^L$ coincides with L. The existence of such an S is guaranteed by Theorem 3.1. This proof (the last part of which may be ignored here) yields subgroups X and Y of G (there denoted $G \cap D$ and V, respectively) that satisfy the conditions of Lemma 3.2 and whose intersection is precisely L. Hence we conclude from Lemma 3.2 that $L\backslash G$ has a representative algebraic structure, so that Theorem 3.3 is proved.

4. Lifting of structures.

We begin with the following almost evident generalization of Theorem 3.4 of [5].

THEOREM 4.1. *Let L be a reductive closed complex Lie subgroup of the complex analytic linear group G. Then every representative algebraic structure of $L\backslash G$ is affine.*

Proof. By Theorem 2.5, there is an isomorphism of algebraic varieties $L\backslash G \rightarrow K\backslash A_S$, where K is an algebraic subgroup of a suitable A_S. Now it follows exactly as in the proof of Theorem 3.4 of [5] that K is a fully reducible algebraic linear group. Hence, from the theory of algebraic homogeneous spaces (see the reference in the proof of Theorm 3.2 of [5]), $K\backslash A_S$ is an affine algebraic variety, which proves Theorem 4.1.

We shall say that a representative algebraic structure on $L\backslash G$ is *liftable* to G if there is a representative algebraic structure on G (which, by Theorem

4.1, is necessarily affine) such that the canonical map $G \rightarrow L\backslash G$ is a morphism of algebraic varieties.

THEOREM 4.2. *Let L be a closed complex Lie subgroup of the complex analytic linear group G, and suppose that $L\backslash G$ has a representative algebraic structure with field F. Then this structure is liftable to G if and only if $A^{F'}$ is a semidirect product $X \cdot L$, where X is an abelian pro-algebraic subgroup of A all whose elements are semisimple automorphisms of $\boldsymbol{R}(G)$.*

Proof. By virtue of Theorem 4.1, the proof is practically identical with the proof of Theorem 5.1 in [5]; one merely has to replace the algebra P of polynomial functions figuring there with F.

THEOREM 4.3. *Let L be a closed complex Lie subgroup of the complex analytic linear group G. If $L\backslash G$ has a representative algebraic structure that is liftable to G then the component group of L is finite. Conversely, if the component group of L is finite then every representative algebraic structure of $L\backslash G$ is liftable to G.*

Proof. Mutatis mutandis, the proof can be read off from the proof of Theorem 5.3 in [5].

This proof actually yields the following results. If L has only a finite number of connected components and $L\backslash G$ has a representative algebraic structure with field $F \subset {}^{L}[\boldsymbol{R}(G)]$ then there is a finitely generated fully stable subalgebra S of $\boldsymbol{R}(G)$ and a complex toroid $X(S)$ in $(A^{F'})_S$ such that the following conditions are satisfied: (1) F and F' are contained in $[S]$; (2) S separates the points of G and contains $\mathrm{Hom}(G, C)$; (3) A_S is the semidirect product $X(S) \cdot G_S$.

Then we have $(A^{F'})_S = (A_S)^{F'} = X(S) \cdot L_S$, the given representative algebraic structure on $L\backslash G$ is that obtained from the identification of $L\backslash G$ with $(X(S) \cdot L_S)\backslash A_S$, and this lifts to the affine representative algebraic structure on G obtained from the identification of G with $X(S)\backslash A_S$.

5. The isomorphism theorem. We consider the situation described just above. With evident slight adjustments in the auxiliary definitions, the arguments given in Section 2 of [6] apply here and yield the following results. Let X be the full inverse image of $X(S)$ in A. Then X is a complex pro-toroid and actually lies in the normalizer $N_A(L)$ of L in A. There is a maximal complex pro-toroid T in $N_A(L)$ such that $X \subset T$. Moreover, A is the semidirect product $X \cdot G$, the complex pro-toroid T is the direct product $X \times (T \cap G)$, and the pro-algebraic group hull L^* of L in A is contained in $X \cdot L$. We shall describe this situation by saying that the given representative algebraic structure on $L\backslash G$ is *associated* with the maximal complex pro-toroid T in $N_A(L)$. From the definition of X, and the fact that $(A^{F'})_S = X_S \cdot L_S$, it is easily seen that $A^{F'} = X \cdot L$. By Proposition 2.4, $F' = [\mathbf{R}(G)]^{X\cdot L}$.

Now let us consider any other representative algebraic structure on $L\backslash G$. This structure is associated with some maximal complex pro-toroid T^0 in $N_A(L)$. It has been shown in Section 2 of [6] that any two maximal complex pro-toroids in $N_A(L)$ are conjugate by an inner automorphism of A effected by an element of $N_G(L)$. Thus there is an element g in $N_G(L)$ such that $gT^0g^{-1} = T$. Our structure on $L\backslash G$ is that obtained by transporting the natural algebraic variety structure of an $(X^0{}_{S^0} \cdot L_{S^0})\backslash A_{S^0}$, where S^0 and X^0 are like S and X above. The conjugation $x \rightarrow gxg^{-1}$ effected by g on G induces an analytic manifold automorphism g^* on $L\backslash G$, because g normalizes L. By enlarging S and S^0, if necessary, we may arrange to have $S^0 = S$. Now consider the following commutative diagram, where the vertical maps are the canonical bijections and the bottom horizontal map is the map induced by the conjugation effected by g_S on A_S:

$$\begin{tikzcd}
L\backslash G \arrow[r, "g^*"'] \arrow[d] & L\backslash G \arrow[d] \\
(X^0{}_S \cdot L_S)\backslash A_S \arrow[r, "g_S{}^*"'] & ((gX^0g^{-1})_S \cdot L_S)\backslash A_S
\end{tikzcd}$$

Clearly, $g_S{}^*$ is an isomorphism of algebraic varieties. The vertical map on the left is an isomorphism of algebraic varieties, with respect to the algebraic structure on $L\backslash G$ that is under consideration. The vertical map on the right is bijective and transports the natural alegbraic variety structure of $((gX^0g^{-1})_S \cdot L_S)\backslash A_S$ to a representative algebraic structure on $L\backslash G$ that is associated with $gT^0g^{-1} = T$. It is clear from the diagram that g^* is therefore an isomorphism of algebraic varieties from $L\backslash G$ with the given algebraic structure to $L\backslash G$ with an algebraic structure that is associated with T.

We shall say that two representative algebraic structures on $L\backslash G$ that are transforms of each other by a map g^* with g in $N_G(L)$ are *G-conjugate.* Notice that g^* leaves the distinguished base point L of $L\backslash G$ fixed, but twists the canonical G-action by the inner automorphism effected by g. However, if one follows up g^* with the canonical translation effected by g on $L\backslash G$ one still has an isomorphism of algebraic varieties, which maps the base point L onto the point Lg, but which commutes with the canonical G-translations on $L\backslash G$. Clearly, everything we said above about X holds also for gX^0g^{-1}.

Conversely, suppose that Y is any complex pro-toroid in A such that

$T = Y \times (T \cap G)$ and $L^* \subset YL$. Then we have $YL = YL^*$, whence it is clear that $(YL)_S$ is an algebraic subgroup of A_S for every finitely generated fully stable subalgebra S of $\boldsymbol{R}(G)$. Exactly as in the proof of Proposition 2.1 in [6], one shows that $Y_S \cap G_S = (1)$ for all sufficiently large such algebras S. This implies that $(YL)_S \cap G_S = L_S$. If, moreover, S separates the points of G then the canonical map $L\backslash G \rightarrow (Y_S \cdot L_S)\backslash A_S$ is bijective and thus defines a representative algebraic structure on $L\backslash G$, which is associated with T exactly as the structure considered at the beginning of this section is associated with T, with Y taking the place of X. If S_1 is any finitely generated fully stable subalgebra of $\boldsymbol{R}(G)$ containing S then the restriction epimorphism $A_{S_1} \rightarrow A_S$ induces an isomorphism of algebraic varieties $(Y_{S_1} \cdot L_{S_1})\backslash A_{S_1} \rightarrow Y_S \cdot L_S)\backslash A_S$, whence it is clear that our representative algebraic structure on $L\backslash G$ is independent of the particular choice of S, and thus is uniquely determined by Y; its field is ${}^{Y \cdot L}[\boldsymbol{R}(G)] = ([\boldsymbol{R}(G)]^{Y \cdot L})'$.

It is evident from the above that *the results of Sections* 3 *and* 4 *of* [6] *concerning the parameterization of quasi-affine structures and their conjugacy under inner or general complex analytic automorphisms of G carry over completely to general representative algebraic structures.* We use the main tool employed there in proving the following isomorphism theorem, which, even in the special case of quasi-affine structures, says more than the statement made at the end of Section 3 in [6].

THEOREM 5.1. *Let G be a complex analytic linear group, and let L be a closed complex Lie subgroup of G. Suppose there are given two representative algebraic structures on $L\backslash G$, and denote the corresponding algebraic varieties by $(L\backslash G)_1$ and $(L\backslash G)_2$. Then there is an algebraic variety isomorphism $(L\backslash G)_1 \rightarrow (L\backslash G)_2$.*

Proof. First, let us assume that L has only a finite number of connected components. In that case, by what we have shown above, we may assume that the given representative algebraic structures are both associated with the same maximal complex pro-toroid T in $N_A(L)$. Let X and Y be complex protoroids of the kind considered above that determine the given algebraic structures on $L\backslash G$. We have $X \times (T \cap G) = T = Y \times (T \cap G)$. There is a finitely generated fully stable subalgebra S of $\boldsymbol{R}(G)$ that satisfies the conditions given at the end of Section 4 with respect to each of our two algebraic structures; in particular, with either X_S or Y_S in the place of $X(S)$. Now the complex toroid T_S has the two decompositions $T_S = X_S \times T \cap G)_S$ and $T_S = Y_S \times (T \cap G)_S$ into direct products of complex toroids. These define a rational homomorphism $\tau \colon X_S \rightarrow (T \cap G)_S$ such that, for each element x of

X_S, $x\tau(x)$ is the component of x in Y_S with respect to the decomposition $T_S = Y_S \times (T \cap G)_S$.

Let U_S denote the maximum unipotent normal subgroup of A_S. It is shown in Section 3 of [6] (in the part preceding the statement of Proposition 3.1) that there is a maximal reductive subgroup H of G such that the following conditions are satisfied: $T \cap G \subset H$, X_S normalizes H_S and $X_S \cdot H_S$ is a maximal fully reducible subgroup of A_S, so that A_S is the composite semidirect product $(X_S \cdot H_S) \cdot U_S$ in the sense of algebraic linear groups; if L^* denotes the pro-algebraic group hull of L in A then

$$L^*_S = (X \cap L^*)_S \cdot ((H_S U_S) \cap L^*_S).$$

Moreover, it is shown there that τ sends $(X \cap L^*)_S$ into L_S.

The elements of A_S may be written uniquely as products xhu, where $x \in X_S$, $h \in H_S$ and $u \in U_S$. Define the map $\sigma: A_S \to A_S$ by $\sigma(xhu) = x\tau(x)hu$. Since the projections of A_S onto X_S, H_S and U_S that correspond to the above decomposition $A_S = (X_S \cdot H_S) \cdot U_S$ are morphisms of algebraic varieties and since τ is also a morphism, it is clear that σ is a morphism of algebraic varieties. We claim that σ sends each coset $X_S L_S p$, where $p \in A_S$, into the coset $Y_S L_S \sigma(p)$. Evidently, it suffices to prove this in the case where p belongs to X_S, which we shall now suppose to be the case. By the above decomposition of L^*_S, every element t of L_S may be written as a product $x_t z_t$ with x_t in $(X \cap L^*)_S$ and z_t in $H_S U_S$. Let x be any element of X_S. Then we have $xtp = xx_t z_t p = (xx_t p)(p^{-1} z_t p)$, and $p^{-1} z_t p \in H_S U_S$. Hence

$$\begin{aligned}\sigma(xtp) &= xx_t p\tau(xx_t p)p^{-1}z_t p = x\tau(xx_t p)x_t z_t p = x\tau(xx_t p)tp \\ &= x\tau(x)\tau(x_t)(\tau(p)t\tau(p)^{-1})\sigma(p).\end{aligned}$$

Since τ sends $(X \cap L^*)_S$ into L_S, we have $\tau(x_t) \in L_S$. Since T_S normalizes L_S, we have $\tau(p)t\tau(p)^{-1} \in L_S$. Hence the last product above lies in $Y_S L_S \sigma(p)$, so that our claim is established.

Thus σ induces a morphism of algebraic varieties

$$\sigma^0: (X_S L_S)\backslash A_S \to (Y_S L_S)\backslash A_S.$$

If we interchange the roles of X_S and Y_S in the above construction of σ^0 we obtain a morphism $(Y_S L_S)\backslash A_S \to (X_S L_S)\backslash A_S$ that inverts σ^0; one must merely observe that H can serve with Y as it did with X. Hence we conclude that σ^0 is actually an isomorphism of algebraic varieties. Since $(L\backslash G)_1$ is canonically isomorphic with $(X_S L_S)\backslash A_S$ and $(L\backslash G)_2$ is canonically isomorphic with $(Y_S L_S)\backslash A_S$, this establishes Theorem 5.1 in the case where L has only a finite number of connected components.

We reduce the general case to the case just disposed of as follows.

There is a finitely generated fully stable subalgebra S of $\boldsymbol{R}(G)$ that separates the points of G and is such that A_S has algebraic subgroups K_1 and K_2 realizing the given representative algebraic structures on $L\backslash G$, so that $(L\backslash G)_i$ is canonically isomorphic with $K_i\backslash A_S$ $(i=1,2)$. Let L_S^* denote the algebraic group hull of L_S in A_S. Then $L_S^* \subset K_1 \cap K_2$ and $L_S^* \cap G_S = L_S$. Hence, for $i=1$ or 2, we have the following commutative diagram of natural maps:

$$\begin{tikzcd}
(L\backslash G)_i \arrow[d] \arrow[dr] & \\
K_i\backslash A_S & L_S^*\backslash(L_S^* G_S) \arrow[l]
\end{tikzcd}$$

If the subscript 1 denotes the connected component of the identity, we have $(L_S^*)_1 \subset (L_S^*)_1 L_S \subset L_S^*$, which shows that $(L_S^*)_1 L_S$ is an algebraic subgroup of L_S^*, whence actually $(L_S^*)_1 L_S = L_S^*$. Hence $L_S^* G_S = (L_S^*)_1 G_S$, so that it is a complex analytic subgroup of A_S. In the diagram above, the vertical map and the slanted map are bijective. Hence the horizontal map is bijective and transports the natural algebraic variety structure of $K_i\backslash A_S$ onto an algebraic variety structure of $L_S^*\backslash(L_S^* G_S)$. Since A_S is an algebraic group hull of $L_S^* G_S$, we know from Theorem 2.2 that this is a representative algebraic structure. Let $(L_S^*\backslash(L_S^* G_S))_i$ denote the algebraic variety with this structure. Since L_S^* is an algebraic linear group, it has only a finite number of connected components. Hence we know from what has already been proved that there is an isomorphism of algebraic varieties of $(L_S^*\backslash(L_S^* G_S))_1$ onto $(L_S^*\backslash(L_S^* G_S))_2$. Since the canonical maps $(L\backslash G)_i \to (L_S^*\backslash(L_S^* G_S))_i$ are isomorphisms of algebraic varieties, this establishes Theorem 5.1 in the general case.

Finally, we remark, without giving the details, that Proposition 1.2 can be used for calculating explicitly the 'deformations' of the fields of representative algebraic structures, as they occurred in our proof of Theorem 5.1, the result being the appropriate elaboration of Theorem 3.2 of [6].

Added in proof. Let G be a complex analytic linear group, and let L be a closed complex Lie subgroup of G such that $L\backslash G$ has a representative algebraic structure, with field $F \subset {}^{L}[\boldsymbol{R}(G)]$. Let u be an element of F, and let u^* denote the corresponding rational function on the algebraic variety $L\backslash G$. On the other hand, let u' denote the meromorphic function on the complex analytic variety $L\backslash G$ that is determined by u. We claim that u^* and u' actually coincide, i. e., that the domain D_{u^*} of the rational function u^* coincides with the domain $D_{u'}$ of the meromorphic function u'.

Evidently, we have $D_{u^*} \subset D_{u'}$. Conversely, let p be a point in $D_{u'}$. Let Q denote the local ring of the rational functions defined at p, and let P

denote the maximum ideal of Q, consisting of the elements of Q that take the value 0 at p. Because of the homogeneity, every point of the algebraic variety $L\backslash G$ is a simple point. Hence Q is a regular local ring and, if n is the dimension of $L\backslash G$, there is an ideal basis $(x_1,\cdots,x_n)$ of P such that the completion Q^0 of Q with respect to P is the ring of formal integral power series $C[[x_1,\cdots,x_n]]$. The ring H of the germs of holomorphic functions at p, belonging to the structure of $L\backslash G$ as a complex analytic variety, may be identified with a subring of Q^0 containing Q, namely with the ring of the convergent power series.

We may write $u^* = a/b$ with a and b in Q. The fact that p belongs to $D_{u'}$ implies that there is an element h in H such that $hb = a$. In particular, this shows that a belongs to Q^0b. By a general result in local algebra [Zariski-Samuel, *Commutative Algebra*, vol. II, p. 262, Th. 9], we have $Q^0b \cap Q = Qb$. Hence a belongs to Qb, i.e., u^* belongs to Q, which means that p lies in D_{u^*}. Thus $D_{u^*} = D_{u'}$, and so $u^* = u'$.

Let π denote the canonical map $G \to L\backslash G$. Then the domain D_u of u in G, as we defined it at the beginning of Section 2, is $\pi^{-1}(D_{u^*})$. On the other hand, it is easy to see that the domain of u as a meromorphic function on G is $\pi^{-1}(D_{u'})$. Hence our above result implies that the domain of u as a meromorphic function on G coincides with D_u as defined in Section 2. In general, this is not true for an arbitrary element of $[\boldsymbol{R}(G)]$ that does not belong to a representative algebraic structure.

REFERENCES.

[1] A. Bialynicki-Birula, G. Hochschild and G. D. Mostow, "Extensions of representations of algebraic linear groups," *American Journal of Mathematics*, vol. 85 (1963), pp. 131-144.

[2] C. Chevalley, *Fondements de la géométrie algébrique*, Faculté des Sciences de Paris, 1958 (mimeographed).

[3] ———, *Classification des groupes de Lie algébriques*, Séminaire, Ecole Normale Supérieur, 1956-58, Exposé 8 (mimeographed).

[4] W. L. Chow, *On the projective embedding of homogeneous varieties*, Symposium in honor of S. Lefschetz, Princeton University Press, 1957, pp. 122-128.

[5] G. Hochschild and G. D. Mostow, "Affine embeddings of complex analytic homogeneous spaces," *American Journal of Mathematics*, vol. 87 (1965), pp. 807-839.

[6] ———, "Deformations of affine embeddings of complex analytic homogeneous spaces, *American Journal of Mathematics*, vol. —.

A REPRODUCING FORMULA FOR INTEGRABLE AUTOMORPHIC FORMS.*

By Clifford J. Earle.

1. Introduction. Let G be a Fuchsian group (perhaps not finitely generated) acting on the unit disk $\Delta=\{z:|z|<1\}$. Fix an integer $q\geqq 2$. We call the holomorphic function ϕ on Δ an *automorphic form* of weight $(-2q)$ if $\phi(Az)A'(z)^q=\phi(z)$ for all z in Δ and A in G.

Let $\omega\subset\Delta$ be a fundamental region for G whose boundary has zero area. Let ϕ be an automorphic form of weight $(-2q)$. Following Bers [3], we say that ϕ belongs to the Banach space $\boldsymbol{A}_q(G)$ of *integrable* forms if $\phi(z)(1-|z|^2)^{q-2}\in L_1(\omega)$ with respect to Lebesgue measure on ω. If $\phi(z)(1-|z|^2)^q\in L_\infty(\Delta)$ we say that ϕ belongs to the space $\boldsymbol{B}_q(G)$ of *bounded* forms. When G is the trivial group we write $\boldsymbol{A}_q$ and $\boldsymbol{B}_q$ for $\boldsymbol{A}_q(G)$ and $\boldsymbol{B}_q(G)$.

The space $\boldsymbol{A}_q$, of course, is defined for real values of $q>1$ and has been considered, for instance, by Innis [5]. We reserve the letter q for integers, using t for arbitrary reals. Our main result is

Theorem 1. *$\boldsymbol{A}_q(G)\subset\boldsymbol{A}_{q+t}$, with a continuous inclusion map, for all $t>1$. Hence, each ϕ in $\boldsymbol{A}_q(G)$ satisfies the reproducing formula*

$$\phi(\zeta)=\frac{2q-1}{\pi}\int\int_\Delta\phi(z)(1-|z|^2)^{2q-2}(1-\bar{z}\zeta)^{-2q}\,dxdy. \tag{1}$$

The second statement follows from the first because (1) holds for all ϕ in $\boldsymbol{A}_{2q}$ ([5], Theorem 1). We prove Theorem 1 in § 2. As an application, in § 3 we prove Bers' theorem that each ϕ in $\boldsymbol{A}_q(G)$ is a Poincaré series ([3], Theorem 2). In § 4 we obtain another reproducing formula for ϕ in $\boldsymbol{A}_q(G)$ closely related to (1). Finally, in § 5 we indicate how to phrase our results for discontinuous groups acting on a domain D conformally equivalent to Δ.

Received April 27, 1965.

Revised June 1, 1965.

* This research was supported by the Army Research Office (DURHAM) through grant number DA-ARO(D)-31-124-G455.

2. Proof of Theorem 1. The identity

$$(2)\qquad \frac{\pi}{t-1}(1-|z|^2)^{-t}=\int\int_\Delta (1-|\zeta|^2)^{t-2}|1-\bar{z}\zeta|^{-2t}\,d\xi d\eta$$

is easily verified for $t>1$ by the change of variables $\zeta=(w+z)/(1+\bar{z}w)$. From (2) we obtain, for ϕ in $A_q(G)$,

$$\begin{aligned}\frac{\pi}{t-1}\|\phi\|_{A_q(G)}&=\frac{\pi}{t-1}\int\int_\omega |\phi(z)|\,(1-|z|^2)^{q-2}\,dxdy\\ &=\int\int_\omega |\phi(z)|(1-|z|^2)^{q+t-2}\,dxdy\int\int_\Delta (1-|\zeta|^2)^{t-2}|1-\bar{z}\zeta|^{-2t}\,d\xi d\eta\\ &=\int\int_\Delta |\phi(z)|(1-|z|^2)^{q+t-2}\,dxdy\int\int_\omega (1-|\zeta|^2)^{t-2}|1-\bar{z}\zeta|^{-2t}\,d\xi d\eta\\ &\geqq 4^{-t}\|\phi\|_{A_{q+t}}\int\int_\omega (1-|\zeta|^2)^{t-2}\,d\xi d\eta.\end{aligned}$$

Since the integral over ω is positive, the theorem is proved.

The interchange of the domains ω and Δ in the above argument is justified by subdividing Δ into images of ω and changing variables of integration. This device has been used by Ahlfors ([1], p. 421) in a similar context. The author is indebted to Professor Ahlfors for the observation that Theorem 1 can also be deduced from the fact that

$$(3)\qquad \sum_{A\in G}(1-|Az|^2)^t \text{ is bounded}$$

for every $t>1$ by a number depending only on t and G. (3) is a special case of Godemont's theorem on the boundedness of Poincaré series ([4], Theorem 5 bis). Ahlfors [2] has given an elementary proof of (3) for $t=2$.

Theorem 1 has the following

COROLLARY. *For each q, $t>1$ there is a constant C, depending only on q, t and the group G, such that*

$$\sup|\phi(z)|(1-|z|^2)^{q+t}\leqq C\|\phi\|_{A_q(G)} \textit{ for all } \phi \textit{ in } A_q(G).$$

Since the left hand side of this inequality is the B_{q+t} norm of ϕ, the corollary is an immediate consequence of Theorem 1 and the fact that $A_{q+t}\subset B_{q+t}$ with a continuous inclusion map ([3], p. 199).

3. Poincaré series. For Φ in A_q the *Poincaré series*

$$\Theta_q\Phi(z)=\sum_{A\in G}\Phi(Az)A'(z)^q$$

converges to an element of $A_q(G)$. Conversely, we shall prove

THEOREM 2 (Bers). *Suppose* $\phi \in A_q(G)$. *Then*

$$\Phi(\zeta) = \frac{2q-1}{\pi} \int\int_{\omega} \phi(z)(1-|z|^2)^{2q-2}(1-\bar{z}\zeta)^{-2q}\, dxdy$$

belongs to A_q, *and* $\Theta_q \Phi = \phi$.

Proof. Using the identity (2) with $t = q$, we conclude that $\Phi \in A_q$. Furthermore,

$$\begin{aligned}\Theta_q\Phi(\zeta) &= \frac{2q-1}{\pi} \sum_{A \in G} A'(\zeta)^q \int\int_{\omega} \phi(z)(1-|z|^2)^{2q-2}(1-\bar{z}A\zeta)^{-2q}\, dxdy \\ &= \frac{2q-1}{\pi} \sum_{A \in G} A'(\zeta)^q \int\int_{A^{-1}(\omega)} \phi(Az)(1-|Az|^2)^{2q-2} \\ &\qquad\qquad \times (1-\overline{Az}A\zeta)^{-2q}\,|A'(z)|^2\, dxdy \\ &= \frac{2q-1}{\pi} \sum_{A \in G} \int\int_{A^{-1}(\omega)} \phi(z)(1-|z|^2)^{2q-2}(1-\bar{z}\zeta)^{-2q}\, dxdy \\ &= \frac{2q-1}{\pi} \int\int_{\Delta} \phi(z)(1-|z|^2)^{2q-2}(1-\bar{z}\zeta)^{-2q}\, dxdy = \phi(\zeta)\end{aligned}$$

by the dominated convergence theorem and Theorem 1. QED

4. **A reproducing formula.** For any ζ in Δ the function

$$\Phi_\zeta(z) = \pi^{-1}(2q-1)(1-z\bar{\zeta})^{-2q}$$

belongs to A_q. Hence its Poincaré series converges and we may define $\alpha_q(z,\zeta) = \Theta_q\Phi_\zeta(z)$. This function has the reproducing property

$$\phi(\zeta) = \int\int_{\omega} \phi(z)\alpha_q(z,\zeta)^*(1-|z|^2)^{2q-2}\, dxdy \tag{4}$$

for ϕ in $A_q(G)$. Indeed, by (1),

$$\begin{aligned}\phi(\zeta) &= \int\int_{\Delta} \phi(z)\Phi_\zeta(z)^*(1-|z|^2)^{2q-2}\, dxdy \\ &= \sum_{A \in G} \int\int_{A(\omega)} \phi(z)\Phi_\zeta(z)^*(1-|z|^2)^{2q-2}\, dxdy \\ &= \sum_{A \in G} \int\int_{\omega} \phi(z)\Phi_\zeta(Az)^*(A'(z)^q)^*(1-|z|^2)^{2q-2}\, dxdy \\ &= \int\int_{\omega} \phi(z)\alpha_q(z,\zeta)^*(1-|z|^2)^{2q-2}\, dxdy.\end{aligned}$$

For ϕ in $A_q(G)$ and ψ in $B_q(G)$ the *Petersson scalar product* (ϕ,ψ) is defined by

$$(\phi,\psi) = \int\int_{\omega} \phi(z)\psi(z)^*(1-|z|^2)^{2q-2}\, dxdy.$$

According to Theorem 1 of [3], every bounded linear function on $\boldsymbol{A}_q(G)$ can be written uniquely in the form (ϕ, ψ), ψ in $\boldsymbol{B}_q(G)$. In particular, for each ζ in Δ there is a unique function ψ_ζ in $\boldsymbol{B}_q(G)$ such that

$$\phi(\zeta) = (\phi, \psi_\zeta) \text{ for all } \phi \text{ in } \boldsymbol{A}_q(G). \tag{5}$$

THEOREM 3. *The unique function* ψ_ζ *in* $\boldsymbol{B}_q(G)$ *with the reproducing property* (5) *is* $\psi_\zeta(z) = \alpha_q(z, \zeta)$.

Proof. Because of (4), we need only prove that $\alpha_q(z, \zeta) \in \boldsymbol{B}_q(G)$ for each fixed ζ in Δ. This follows readily from (3). Indeed,

$$(1 - |z|^2)^q |\alpha_q(z, \zeta)| \leqq \pi^{-1}(2q - 1)(1 - |\zeta|)^{-2q} \sum_{A \in G} (1 - |Az|^2)^q$$

which is a bounded function of z for each fixed ζ in Δ. QED

5. Extensions to other domains. In [3], Bers considers spaces $\boldsymbol{A}_q(D, G)$ and $\boldsymbol{B}_q(D, G)$ where D is conformally equivalent to Δ and G is a discontinuous group of automorphisms of D. Briefly, let $f \colon \Delta \to D$ be conformal. Then $G' = f^{-1} \circ G \circ f$ is a Fuchsian group, and the map $\phi \to (\phi \circ f)(f')^q$ defines isometries of $\boldsymbol{A}_q(D, G)$ on $\boldsymbol{A}_q(G')$ and $\mathbf{B}_q(D, G)$ on $\boldsymbol{B}_q(G')$. Once again we use notations $\boldsymbol{A}_q(D)$ and $\boldsymbol{B}_q(D)$ when G is the trivial group.

The formulas (1) and (4) have obvious extensions for ϕ in $\boldsymbol{A}_q(D, G)$ or $\boldsymbol{B}_q(D, G)$, obtained by the isometry and a change of variables. However, the inclusion $\boldsymbol{A}_q(G) \subset \boldsymbol{A}_{q+t}$ depends on non-invariant properties of the disk. In general, let $k(z, \zeta)$ be the Bergman kernel function of D. Then for each ζ in D and each $t > 1$, the map $\phi(z) \to \phi(z) k(z, \zeta)^t$ is a continuous mapping of $\boldsymbol{A}_q(D, G)$ into $\boldsymbol{A}_{q+t}(D)$. The proof is immediate, using the isometry.

INSTITUTE FOT ADVANCED STUDY.

REFERENCES.

[1] L. V. Ahlfors, "Finitely generated Kleinian groups," *American Journal of Mathematics*, vol. 86 (1964), pp. 413-429.

[2] ———, "Eine Bemerkung über Fuchssche Gruppen," *Mathematische Zeitschrift*, vol. 84 (1964), pp. 244-245.

[3] L. Bers, "Automorphic forms and Poincaré series for infinitely generated Fuchsian groups," *American Journal of Mathematics*, vol. 87 (1965), pp. 196-214.

[4] R. Godement, "Séries de Poincaré et Spitzenformen," *Séminaire Henri Cartan*, 10e année (1957/1958), exposé 10.

[5] G. S. Innis. Jr., "Some reproducing kernels for the unit disk," *Pacific Journal of Mathematics*, vol. 14 (1964), pp. 177-187.

ON EXTENSIONS OF LOCALLY COMPACT GROUPS.

By MARC A. RIEFFEL.*

Let Q be a locally compact group, and let K be a Q-module, that is, a locally compact Abelian group on which Q operates continuously as a group of automorphisms. An extension of Q by K is a locally compact group, G, together with a homomorphism, p, of G onto Q, and a topological isomorphism, i, of K onto the kernel of p, such that p induces a topological isomorphism of $G/i(K)$ onto Q in such a way that the natural action of $Q=G/i(K)$ on $K=i(K)$ by inner automorphisms is the given action of Q on K. (See [8] for an extensive bibliography).

Mackey [6] showed that if K and Q are separable, in the sense of satisfying the second axiom of countability, then to each extension, G, of Q by K there corresponds a (non-unique) normalized Borel 2-cocycle on Q with values in K, that is, a Borel function c from $Q\times Q$ into K satisfying the "cocycle identity"

$$c(u,vw)\,[c(v,w)]^{u}=c(uv,w)\,c(u,v) \qquad u,v,w\in Q$$

and the normalization

$$c(u,v)=e_K \text{ whenever } u \text{ or } v \text{ is } e_Q$$

(where s^u denotes the image of $s\in K$ under the automorphism of K corresponding to $u\in Q$). Conversely, Mackey showed how to construct for each normalized Borel 2-cocycle a corresponding extension of Q by K, in such a way that the extensions of Q by K are classified by certain equivalence classes, the cohomology classes, of 2-cocycles, in the manner first described by Schreier for finite groups.

The purpose of this note is to give a new discussion of the construction of an extension corresponding to a given 2-cocycle which we feel is simpler than Mackey's original discussion in that it does not use the theory of standard Borel structures. This also permits us to weaken considerably the measurability conditions imposed. In particular Q can be an arbitrary locally

---

Received September 13, 1965.

* This research was partially supported by National Science Foundation grant GP-4020.

compact group, and K is only required to be σ-compact. (At the end of this note we give an example to show that this condition is necessary in order for our approach to work.)

The question of whether for a given extension there always exist corresponding cocycles satisfying any useful measurability condition is, as far as we know, still quite unanswered in the non-separable case.

We are indebted to Calvin C. Moore for having brought this problem to our attention.

1. The measurability conditions. In this section we describe and motivate the measurability conditions which we will impose on the cocycles under consideration. We use the following notation. If H is any locally compact group, then $B(H)$ will denote the σ-ring of Baire subsets [2] of H, and $B_l(H)$ will denote the σ-ring of locally Baire subsets, that is, subsets whose intersection with every Baire set is a Baire set.

If (G, i, p) is an extension of Q by K, then the usual way of associating a cocycle (as we will abbreviate "normalized 2-cocycle" from now on) to an extension is by constructing a cross-section of p, that is, a function, q, from Q into G satisfying $pq = id_Q$ and normalized so that $q(e_Q) = e_G$. Then the cocycle, c, corresponding to q is defined by

$$c(u, v) = q(u)q(v)[q(uv)]^{-1}.$$

It is evident that the range of c is in the kernel of p, and it follows easily from the associative law in G that c is a cocycle.

If q satisfies certain measurability conditions, then so will c. For instance, Mackey showed [5] that if K and Q are separable, then a cross-section, q, can always be constructed which is a Borel function, and in this case c will be a Borel function. In the case of arbitrary locally compact groups we do not know how, in general, to construct cross-sections satisfying any useful measurability conditions. However, a natural condition to impose on cross-sections is that if $E \in B(G)$ then $q^{-1}(E) \in B(Q)$. In this case, as long as K is σ-compact, the measurability condition satisfied by c is given by

PROPOSITION 1. *If K is σ-compact, if q is a cross-section such that $q^{-1}(E) \in B(Q)$ whenever $E \in B(G)$, and if c is the corresponding cocycle, then $c^{-1}(A) \in B_l(Q \times Q)$ for all $A \in B(K)$.*

Before we give the proof it will be useful to have

PROPOSITION 2. *Let G be a locally compact group, and let K be a closed normal subgroup of G such that K and G/K are σ-compact. Then G is σ-compact.*

Proof. Let U be a compact symmetric neighborhood of e_G, and let V_n be an increasing sequence of compact symmetric subsets of K whose union is K. Let $G_0 = \bigcup_{n=1}^{\infty} (U \cup V_n)^n$. Then G_0 is an open σ-compact subgroup of G which contains K. To conclude the proof it suffices to show that the discrete coset space G/G_0 is countable. Now G_0/K is an open subgroup of G/K, and since G/K is σ-compact, the coset space $(G/K)/(G_0/K)$ is countable. But there is an evident bijection between G/G_0 and $(G/K)/(G_0/K)$.

Proof of Proposition 1. We must show that if $D \in B(Q \times Q)$ then $D \cap c^{-1}(A) \in B(Q \times Q)$ for all $A \in B(K)$. Choose $E \in B(Q)$ such that $D \subseteq E \times E$. We can find an open σ-compact subgroup, Q_0, of Q such that $E \subseteq Q_0$. Then it suffices to show that the restriction of c to $Q_0 \times Q_0$ is a Baire function. Let $G_0 = p^{-1}(Q_0)$. Then G_0 is an extension of Q_0 by K, and so is σ-compact by Proposition 2. Since q is a cross-section, its restriction to Q_0 is a Baire function which has values in G_0. Then the restriction of c to $Q_0 \times Q_0$ is the composition of the maps

$$\begin{aligned}(u,v) &\to (u,v,uv) \to (q(u),q(v),q(uv)) \\ &\to (q(u),q(v),q(u)q(v)[q(uv)]^{-1}) \to q(u)q(v)[q(uv)]^{-1}\end{aligned}$$

each of which is easily seen to be a Baire map, as long as q is considered to have values in Q_0.

Thus one type of measurability condition which we will impose on cocycles is

M$_1$). If $A \in B(K)$ *then* $c^{-1}(A) \in B_l(Q \times Q)$.

There is another type of measurability condition on cocycles for which we can also carry out our construction of an extension. We use the following notation. If H is any locally compact group, then $M(H)$ will denote the σ-ring of all subsets of H which are measurable with respect to Haar measure on H, and $M_l(H)$ will denote the σ-ring of all subsets of H which are locally measurable with respect to $M(H)$. Then our measurability condition is

M$_2$). If $A \in B(K)$ then $c^{-1}(K) \in M_l(Q) \times M_l(Q)$.

Since we have not found any measurability condition on cross-sections which implies that the corresponding cocycles satisfy condition M$_2$, it is not evident whether this condition will ever be useful.

2. Construction of the extension group. In this section K is assumed to be σ-compact. We suppose given a cocycle, c, satisfying either condition

M_1 or M_2. We show how to construct an extension, G, of Q by K corresponding to c.

As usual, G as a set is $K \times Q$, and the law of composition in G is defined by

$$(r, u)(s, v) = (rs^u c(u, v), uv).$$

Then it is not difficult to check that G is a group, with identity element $e_G = (e_K, e_Q)$, and that if an injection $i \colon K \to G$ is defined by $i(r) = (r, e_Q)$, and a projection $p \colon G \to Q$ is defined by $q(r, u) = u$, then G is an algebraic extension of Q by K. The entire difficulty consists of placing a suitable topology on G.

If c were continuous, then the product topology on G would do. But Hochschild showed [4, p. 105] that even for compact Lie groups one cannot always find continuous cocycles corresponding to a given extension. If c is not continuous, then the law of composition in G is not continuous with respect to the product topology. Mackey's ingenious idea was that even though the product topology is not the appropriate topology, nevertheless for separable groups the product measure structure is the appropriate measure structure as long as the cocycle is a Borel function. We show that this is still the case under our more general hypotheses (though our example in Section 4 shows that this need not be the case if K is not σ-compact). Specifically, we show that if m and n denote right Haar measures on K and Q respectively, if $\mu = m \times n$ is the product measure on G, and if c satisfies condition M_1, then μ on the product Baire ring on G will be a right invariant measure (though, surprisingly, the product of left Haar measures will not be left invariant; e.g. let $K = Q =$ real numbers, $c \equiv 0$, $r^u = r(\exp u)$). A similar but slightly weaker result holds for condition M_2. We can then define the Weil topology on G as described by Halmos [2].

We now prove the above assertions. There are two σ-rings on G which will be useful, namely

$$B(G) = B(K) \times B(Q), \text{ and } (B \times M)(G) = B(K) \times M(Q).$$

Similarly there are three σ-rings on $G \times G$ which will be useful, namely

$$\begin{aligned} B(G \times G) &= B(G) \times B(G),\ (B \times M)^2(G \times G) \\ &= (B \times M)(G) \times (B \times M)(G), \text{ and } M(G \times G), \end{aligned}$$

the domain of the completion of $\mu \times \mu$.

As in [2, p. 257] we define bijections of $G \times G$ onto itself as follows

$$
\begin{aligned}
R(x,y) &= (x,yx), & R^{-1}(x,y) &= (x,yx^{-1})) \\
S(x,y) &= (x,xy), & S^{-1}(x,y) &= (x,x^{-1}y) \\
W(x,y) &= (yx,x^{-1}).
\end{aligned}
$$

We note that W is its own inverse. We also denote by I the bijection of G onto itself defined by $I(x) = x^{-1}$. Our first task is to study the measurability of these maps with respect to the σ-rings defined above.

LEMMA A. *Let* $E \in B(G \times G)$ *and* $F \in B(G)$. *If* c *satisfies condition* M_1 *[resp.* M_2*], then* $R(E)$, $R^{-1}(E)$, $S(E)$, $S^{-1}(E)$, $W(E) \in B(G \times G)$ *[resp.* $(B \times M)^2(G \times G)$*], and* $I(F) \in B(G)$ *[resp.* $(B \times M)(G)$*].*

Proof. We give the proof only for R^{-1}, the proof for the other maps being only slightly more complicated.

If $x, y \in G$, $x = (r,u)$, $y = (s,v)$, then

$$R(x,y) = R(r,u,s,v) = (r,u,sr^{v}c(v,u),vu).$$

Thus R is the composition of the following maps, each of which is measurable with respect to the σ-ring indicated, if c satisfies condition M_1.

$$
\begin{aligned}
\underset{B(G\times G)}{(r,u,s,v)} \rightarrow\ & \underset{B(K\times K)\times B_{\iota}(Q\times Q)\times B(Q\times Q)}{(r,s,u,v,u,v)} \\
& \rightarrow \underset{B(K^3\times Q^2)}{(r,s,c(v,u),u,v)} \rightarrow \underset{B(K^3\times Q^2)}{(r,sr^{v},sr^{v}c(v,u),u,vu)} \\
& \rightarrow \underset{B(G\times G)}{(r,u,sr^{v}c(v,u),vu)}.
\end{aligned}
$$

(We note in particular that the measurability of the last map depends on the σ-compactness of K). Thus if $E \in B(G \times G)$, then $(R)^{-1}(E) \in B(G \times G)$. But $(R)^{-1}(E) = R^{-1}(E)$. If c satisfies condition M_2, then the first two σ-rings must be replaced by

$$(B \times M)^2(G \times G) \text{ and } B(K \times K) \times M_{\iota}(Q) \times M_{\iota}(Q) \times B(G \times G)$$

respectively.

The next three lemmas (but not necessarily their proofs) parallel the theorems in Section 59 of [2] which are labeled by the same letter. To begin with, we show that R, R^{-1} and W are measure preserving.

LEMMA B. *If* c *satisfies either condition* M_1 *or* M_2, *and if* $E \in M(G \times G)$, *then* $R(E)$, $R^{-1}(E)$ *and* $W(E) \in M(G \times G)$, *and*

$$\mu \times \mu(R(E)) = \mu \times \mu(R^{-1}(E)) = \mu \times \mu(W(E)) = \mu \times \mu(E).$$

Proof. Suppose first that $E \in B(G \times G)$ so that Lemma A applies. Then

$R^{-1}(E) \in B(G\times G)$ or $(B\times M)^2(G\times G)$, and in either case $\mu\times\mu(R^{-1}(E))$ is defined and Fubini's theorem can be used. Then

$$\begin{aligned}\mu\times\mu(R^{-1}(E)) &= \int\chi_{R^{-1}(E)}(x,y)\,d\mu(x)\,d\mu(y) = \int\chi_E(x,yx)\,d\mu(x)\,d\mu(y)\\ &= \int\int\Big(\int\Big(\int\chi_E(r,u,sr^vc(v,u),vu)\,dm(s)\Big)\,dn(v)\Big)\,dm(r)\,dn(u)\\ &= \mu\times\mu(E)\end{aligned}$$

by the right invariance of m and n. Similarly $\mu\times\mu(R(E)) = \mu\times\mu(E)$ for $E\in B(G\times G)$. In particular, R and R^{-1} carry null sets into null sets. Since R and R^{-1} are inverse to each other, it follows that they are both measure preserving maps of $M(G\times G)$ onto itself. Fnally, $W = R^{-1}TR$ where $T(x,y) = (y,x)$, which is clearly measure preserving.

Lemma C. *For any* $E,F\subseteq G$ *and for any* $x\in F$

$$R(F\times E)_x = Ex,\quad S(F\times E)_x = xE,\quad W(F\times E)_x = Fx^{-1}\cap E^{-1}.$$

We are using Halmos' notation for sections of sets. We omit the verification.

Lemma D. *Let* c *satisfy condition* M_1 [*resp.* M_2]. *If* $E\in B(G)$ *and* $x\in G$ *then* $xE, Ex, E^{-1}\in B(G)$ [*resp.* $(B\times M)(G)$], *and* $\mu(Ex) = \mu(E)$. *If also* $\mu(E) > 0$, *then* $\mu(xE) > 0$ *and* $\mu(E^{-1}) > 0$.

Proof. Choose $F\in B(G)$ such that $x\in F$. Then $F\times E\in B(G\times G)$, so $R(F\times E)\in \mathrm{B}(G\times G)$ [resp. $(B\times M)^2(G\times G)$], so $Ex = R(F\times E)_x \in B(G)$ [resp. $(B\times M)(G)$].

By applying S instead of R we obtain the corresponding fact for xE.

We already know from Lemma A that $E^{-1}\in B(G)$ [resp. $(B\times M)(G)$].

Also, if $x^{-1} = (r,u)$, then

$$\mu(Ex) = \int\chi_E(sr^vc(v,u),vu)\,dm(s)\,dn(v) = \mu(E).$$

Next, if $\mu(E) > 0$, then $\mu\times\mu(W(E\times E)) = \mu\times\mu(E\times E) > 0$, so that for some $x\in G$, $0 < \mu(W(E\times E)_x) = \mu(Ex^{-1}\cap E^{-1})$. Thus $\mu(E^{-1}) > 0$.

Finally, we can find $F\in B(G)$, $F\subseteq E^{-1}$, such that $\mu(F) = \mu(E^{-1})$. Then $\mu(Fx^{-1}) = \mu(F) > 0$. Thus we can find $F_1\in B(G)$, $F_1\subseteq Fx^{-1}$, such that $\mu(F_1) = \mu(Fx^{-1})$. Thus $\mu(F_1^{-1}) > 0$. But $F_1^{-1}\subseteq xF^{-1}\subseteq xE$.

Now if c satisfies condition M_1, the above lemmas are more than enough to prove that $(G, B(G), \mu)$ is a (right) measurable group as defined in [2], and so the Weil topology can be defined on G exactly as described in [2].

If c satisfies condition M_2 then G need not be a measurable group, but the Weil topology can still be defined on G by making only minor changes to the remaining theorems in Sections 59 and 62 of [2]. Our treatment above of Halmos' Theorems B-D of Section 59 should give an indication of how this may be done. For instance, in most places where a set in $(B \times M)(G)$ is obtained, the next move is to replace it by a subset in $B(G)$ of the same measure, as done at the end of the proof of Lemma D above.

The version of Halmos' Theorem E of Section 62 which we obtain is

THEOREM E. *Let c satisfy condition* M_1 *or* M_2. *Let N be the collection of all sets of the form $\{x\colon \mu(E\Delta Ex^{-1}) < \epsilon\}$ where $E \in B(G)$, $0 < \mu(E) < \infty$, and $0 < \epsilon < 2\mu(E)$. Let N' be the collection of all sets of the form EE^{-1} where $E \in B(G)$, $0 < \mu(E) < \infty$. Then N and N' are both neighborhood bases at e_G for the same Hausdorff topology on G, for which G is a topological group.*

The only part of the proof of this theorem which still needs to be checked is that in our situation G is "separated" [2, p. 273], so that the Weil topology is Hausdorff.

LEMMA 1. *G is separated, that is, for any $x \in G$, $x \neq e_G$, there exists $E \in B(G)$, $0 < \mu(E) < \infty$ such that $\mu(E\Delta Ex^{-1}) > 0$.*

Proof. Let $x = (r, u)$.

Case I. $u = e_Q$, so that $r \neq e_K$. Since K is Hausdorff, we can choose $A \in B(K)$, $0 < m(A) < \infty$, such that $m(A\Delta Ar^{-1}) > 0$. Let $C \in B(Q)$ be any neighborhood of e_Q such that $0 < n(C) < \infty$. Let $E = A \times C$. Then

$$\mu(E\Delta Ex^{-1}) = \int |\chi_E(s,v) - \chi_E(sr^v c(v,e), ve)|\, dm(s)\, dn(v)$$

$$= \int_C m(A\Delta A(r^{-1})^v)\, dn(v).$$

Now by [2, Theorem A, p. 266] $m(A\Delta A(r^{-1})^v)$ is a continuous function of v. Since it is non-negative, and is strictly positive at $v = e_Q$, and since C is a neighborhood of e_Q, it follows that $\mu(E\Delta Ex^{-1}) > 0$.

Case II. $u \neq e_Q$. Since Q is Hausdorff we can choose $C \in B(Q)$, $0 < n(C) < \infty$, such that $n(C\Delta Cu^{-1}) > 0$. We note that since

$$n(C\Delta Cu^{-1}) + 2n(C \cap Cu^{-1}) = 2n(C),$$

it follows that $n(C - Cu^{-1}) > 0$. Choose $A \in B(K)$, $0 < m(A) < \infty$, and let $E = A \times C$. Then

$$\mu(E\Delta Ex^{-1}) = \int |\chi_E(y) - \chi_E(yx)|\, d\mu(y)$$

$$\geqq \int_{A\times C} [1 - \chi_A(sr^v c(v,u))\chi_C(vu)]\, dm(s)\, dn(v)$$

$$\geqq \int_{A\times C} [1 - \chi_C(vu)]\, dm(s)\, dn(v) = m(A)n(C - Cu^{-1}) > 0.$$

Our definition of extension includes the requirement that the extension group be locally compact. But the Weil topology, while it is locally bounded, need not be locally compact [2, p. 276]. Thus the proof that G is locally compact under our hypotheses will be an important step in our development. Our tactics consist of proving first that G is an extension of Q by K only as topological groups. It will then follow from a theorem on operations in the class of locally compact groups that G is locally compact.

3. The proof that G is an extension of Q by K. We have indicated earlier that (G, i, p) is an algebraic extension of Q by K. We need only verify that i is a topological embedding, that p is continuous, and, most important, that Q has the quotient topology. It is well known that it is sufficient to verify these facts at the identity elements of the groups in question.

LEMMA 2. *The homomorphism i is a topological embedding.*

Proof. This is equivalent to the statement that if r_a is a net of elements in K, then $r_a \rightarrow e_K$ iff $i(r_a) = (r_a, e_Q) \rightarrow e_G$. But this is the statement that $r_a \rightarrow e_K$ iff $\mu(E\Delta E(r_a, e_Q)^{-1}) \rightarrow 0$ for every $E \in B(G)$ with $0 < \mu(E) < \infty$. Now

$$\mu(E\Delta E(r_a, e_Q)^{-1}) = \int |\chi_E(s, v) - \chi_E(s(r_a)^v, v)| \, dm(s) \, dn(v)$$

and the cocycle does not appear in this integral. Thus our desired result can be viewed as just the fact that for the usual semidirect product of Q by K [3, p. 210], the ordinary product topology on $K \times Q$ is the same as the topology defined by the neighborhood base of sets of the form $\{x\colon \mu(E\Delta Ex^{-1}) < \epsilon\}$ for $E \in B(G)$, $0 < \mu(E) < \infty$ (where, of course, the law of composition here is that for the semi-direct product).

We remark that it follows that $i(K)$ is a closed subgroup of G, for it is a locally compact, and so locally closed, subgroup of its closure, and so is an open, hence closed, subgroup of its closure [1, p. 42].

LEMMA 3. *The projection p is continuous.*

Proof. Let $x_a = (r_a, u_a)$ be any net in G which converges to e_G. We must show that $u_a \rightarrow e_Q$. Let C be any neighborhood of e_Q. We must show that u_a is eventualy in C. Let D be a compact Baire symmetric neighborhood of e_Q such that $D^2 \subseteq C$. Choose $A \in B(K)$ such that $0 < m(A) < \infty$, and let $E = A \times D$. Then $\mu(E\Delta Ex_a^{-1}) \rightarrow 0$. But

$$\mu(E\Delta Ex_a^{-1}) \geqq m(A)n(D - Du_a^{-1})$$

as in Lemma 1, Case II. Thus eventually $n(D - Du_a^{-1}) < 2n(D)$, so that $D \cap Du_a^{-1} \neq \emptyset$, so that $u_a \in C$.

LEMMA 4. *Q has the quotient topology.*

Proof. It is sufficient to show that if $C \subseteq Q$, and if $p^{-1}(C)$ is a neighborhood of e_G, then C is a neighborhood of e_Q. To show this we use the fact that the sets of the form EE^{-1}, for $E \in B(G)$, $0 < \mu(E) < \infty$, form a base for the neighborhood system at e_G. If $p^{-1}(C)$ is a neighborhood of e_G, choose $E \in B(G)$, $0 < \mu(E) < \infty$, such that $EE^{-1} \subseteq p^{-1}(C)$. Then $\chi_E \in L^1(G, \mu)$. Define an extended real valued function, f, on Q by

$$f(u) = \int_K \chi_E(r, u)\, dm(r).$$

Then by Fubini's theorem $f \in L^1(Q, n)$, and $\int f\, dn = \mu(E) > 0$. Let $D = \{u : f(u) > 0\}$. Then $D \in B(Q)$ and $n(D) > 0$. Furthermore, if $u \in D$, then $\int_K \chi_E(r, u)\, dm(r) > 0$, so that $u \in p(E)$, that is, $D \subseteq p(E)$. Thus $DD^{-1} \subseteq p(EE^{-1}) \subseteq C$, so that C is a neighborhood of e_Q.

Now that we have shown that G is an extension of Q by K as topological groups, we can apply the theorem [7, p. 52] which states that any such extension is locally compact if both K and Q are. Thus G is locally compact.

We remark that if Q, as well as K, is σ-compact, then so is G by Proposition 2.

4. *Example.* We now give an example to show that we need the hypothesis that K be σ-compact in order for our methods to work. Specifically our example shows that if K is not σ-compact, then the product measure structure on $K \times Q$ may have little relation to the appropriate measure structure on $K \times Q$.

As our example, let $K = R_d$, the discrete real line, which is not σ-compact. Let $Q = R$, the ordinary real line, and let $G = R_d \times R$, the ordinary product group. Instead of the usual cross-section, $u \to (0, u)$, of Q into G, we define the cross-section q by $q(u) = -\frac{1}{2}(u^2, u)$. Although q is not continuous, it is Baire measurable, but for the trivial reason that $q^{-1}(E)$ is a countable set for any $E \in B(G)$. The cocycle corresponding to q is $c(u, v) = uv$. Again, although c is not continuous, it is Baire measurable, but for the same trivial reason.

If we try to recover the extension G from K, Q, and c by the construction described in Section 2, we would define the group $G_0 = R_d \times R$, with the product Baire structure, but with law of composition defined by

$$(r, u) + (s, v) = (r + s + uv, u + v).$$

Then not only is the multiplication not continuous with respect to the product topology, but it is also not measurable with respect to the product measure structure. In fact it is easily checked that any translate of a measurable set of strictly positive measure by any element not in K will be non-measurable, for its projection on K will be uncountable

To state the situation slightly differently, we would expect an isomorphism of G onto G_0 to be defined by $(r,u) \rightarrow (r+\frac{1}{2}u^2, u)$. This is indeed an algebraic isomorphism, but is not measurable.

One might consider removing the condition that K be σ-compact, and strengthening the measurability condition on c by requiring that $c^{-1}(A) \in B_l(Q \times Q)$ for all $A \in B_l(K)$. We have not been able to carry out the construction described in Section 2 in this case.

UNIVERSITY OF CALIFORNIA, BERKELEY.

REFERENCES.

[1] N. Bourbaki, *Éléments de mathématiques*, Livre III, *Topologie général*, Chapitre 1, *Structures topologiques*, Actualitiés Sci. Indust. No. 1142 3e ed. Hermann. Paris, 1961.

[2] P. Halmos, *Measure theory*, Van Nostrand, New York, 1950.

[3] E. Hewitt and K. Ross, *Abstract harmonic analysis*, Academic Press, New York, 1963.

[4] G. Hochschild, "Group extensions of Lie groups I," *Annals of Mathematics*, vol. (2) 54 (1951), pp. 96-109.

[5] G. W. Mackey, "Induced representations of locally compact groups I," *Annals of Mathematics*, vol. (2) 55 (1952), pp. 101-139.

[6] ———, "Les ensembles Boréliens et les extensions des groupes," *Journal de Mathématiques Pures et Appliquées*, vol. 36 (1957), pp. 171-178.

[7] D. Montgomery and L. Zippen, *Topological transformation groups*, Interscience, New York, 1955.

[8] C. C. Moore, "Extensions and low dimensional cohomology theory of locally compact groups I," *Transactions of the American Mathematical Society*, vol. 113 (1964), pp. 40-63.

SPECIAL DIVISORS AND ANALYTIC SUBLOCI OF TEICHMUELLER SPACE.

By HERSHEL M. FARKAS.*

Introduction. This paper concerns itself with the following two problems: what restrictions, if any, are imposed on the moduli of a compact Riemann surface by the existence on the surface of special divisors (equivalently a meromorphic function) of fixed degree $r < g$? What restrictions are imposed on the moduli of a compact Riemann surface by an even theta function vanishing at the origin? We show that the second problem in a special case of the first and prove the often quoted but never proven (at least to my knowledge) theorem that an even theta does not vanish at the origin for every surface. We prove (Section 2, Theorem 1′) that the topological cartesian product, $S_r \times S_r \times T^g(S)$, where S_r is the r-fold symmetric product of the surface with itself and $T^g(S)$ is the Teichmueller space, may be endowed with local coordinates turning it into a non-trivial complex analytic fibre space containing $T^g(S)$ as base manifold. This space will be denoted by $W_{r,r}$. Using this space we prove the following main theorem (Section 2, Theorem 2). If ζ, ω are completely distinct equivalent special divisors of degree $r < g$ on S_0, a compact Riemann surface of genus g, then considering the triple $(\zeta, \omega, S_0(T))$, where $S_0(T)$ is a Teichmueller surface lying over S_0, as a point in $W_{r,r}$, in a sufficiently small neighborhood of $(\zeta, \omega, S_0(T))$ there is a g codimensional submanifold of $W_{r,r}$, containing the point $(\zeta, \omega, S_0(T))$ which projects onto a λ codimensional manifold of the given neighborhood of $S_0(T)\lambda = i(\zeta\omega)$ and bounds are obtained for λ. For the definition of Teichmueller surface the reader is referred to Rauch's paper [14].

We use the following notations:

S—always denotes a compact Riemann surface of genus g
$S(T)$—a convenient Teichmueller surface lying over S
S^r—the r-fold cartesian product of S with itself
S_r—the r-fold symmetric product of S with itself

Revised received July 25, 1966.

* The author wishes to express his gratitude to Professor H. E. Rauch for his advice and encouragement during the preparation of this paper. This paper is part of the author's doctoral dissertation submitted to the Belfer Graduate School of Science, Yeshiva University.

ζ, ω—divisors on S, or integrals of first kind

$\Omega(\zeta)$—space of differentials whose divisors are multiples of ζ

$i(\zeta)$—index of specialty of the divisor ζ or the dimension of $\Omega(\zeta)$

$l(\zeta)$—space of meromorphic functions on S whose divisors are multiples of ζ

$r(\zeta)$—dimension of $l(\zeta)$

$d\zeta_i$—the i-th normal differential of the first kind

$d\eta_{p,q}$—the normal differential of the third kind with poles at p and q with residues ± 1 resp.

C^g—space of g complex variables.

For general references the reader should consult any of the standard texts on Riemann surfaces.

An Abelian differential of the first kind on S is a tensor ϕ satisfying the transformation law $\phi(z) = \phi'(z') \dfrac{dz'}{dz}$ where ϕ is everywhere finite and analytic and where z and z' are the respective local coordinates at points p and p' on S. It is well known that on S there exist g linearly independent Abelian differentials of the first kind over the complex numbers.

We assume that S is endowed with a canonical homotopy and homology basis $\gamma_1, \cdots, \gamma_g, \delta_1, \cdots, \delta_g$ and determine the Riemann matrix by letting

$$\int_{\gamma_j} d\zeta_i = \delta_{ij} \qquad \int_{\delta_j} d\zeta_i = \Pi_{ij}.$$

Actually we may often, by an abuse of language, call the $g \times g$ matrix $\| \Pi_{ij} \|$ the Riemann matrix. It follows immediately from the bilinear relations of Riemann that $\| \Pi_{ij} \|$ is symmetric and that its imaginary part is positive definite.

A divisor on S is a set of points and associated integers $p_1^{\alpha_1} \cdots p_n^{\alpha_n}$. $\sum_{k=1}^{n} \alpha_k = d$ is called the degree of the divisor. Observe that divisors may be multiplied and divided. Multiplication being defined by juxtaposition and division being defined by

$$\frac{p_1^{\alpha_1} \cdots p_n^{\alpha_n}}{Q_1^{B_1} \cdots Q_m^{B_m}} = p_1^{\alpha_1} \cdots p_n^{\alpha_n} Q_1^{-\beta_1} \cdots Q_m^{-\beta_m} \cdots Q_r^{-\beta_r}.$$

Consider now a meromorphic function f on S. Denote the zeros of f by $p_1^{\alpha_1} \cdots p_n^{\alpha_n}$ (the superscripts denoting multiplicities) and the poles of f by $Q_1^{\beta_1} \cdots Q_r^{\beta_r}$. We define the divisor of f to be $p_1^{\alpha_1} \cdots p_n^{\alpha_n} Q_1^{-\beta_1} \cdots Q_r^{-\beta_r}$.

Since a meromorphic function f on S has the same number of zeros as poles, the degree of the divisor of f is zero. The divisor of a differential is defined in an analogous manner. We define an equivalence relation between divisors of the same degree as follows: $\zeta \sim \omega$ iff ζ/ω is the divisor of a meromorphic function on S. A divisor is called integral if $\alpha_k \geqq 0$ for each k. If ζ/ω is integralwe say that ζ is a multiple of ω.

Let ζ be a divisor on S. We denote by $l(\zeta)$ and $\Omega(\zeta)$ the vector spaces of meromorphic functions and differentials of the first kind whose divisors are multiples of ζ. Let $r(\zeta)$ and $i(\zeta)$ be the dimension of $l(\zeta)$ and $\Omega(\zeta)$ respectively. One may then show quite easily that $r(\zeta)$ and $i(\zeta)$ depend only on the divisor class of ζ.

Let $d\zeta_1, \cdots, d\zeta_g$ be a normal basis for the Abelian differentials of the first kind on S. Let P_0 be a fixed point on S and consider the map: $u\colon S \rightarrow C^g$ $u(p) = u_1(p) \cdots u_g(p))$ where $u_i(p) = \int_{p_0}^{p} d\zeta_i$. u_i is a finite everywhere regular multivalued function on S. The multivaluedness arises from the dependence on the path of integration. In order to make u_i single valued we simply identify all possible images of a point p. Since any two images of p can differ only by an integral linear combination of the columns of the Riemann matrix, we identify all such points. The space C^g under this identification is a compact commutative complex Lie group of complex dimension g called the Jacobian variety.

u has been defined as a map from $S \rightarrow J(S)$; however, it can be extended to a map from $S_r \rightarrow J(S)$. The definition is the natural one. We define $u(\zeta)\zeta = (p_1^{\alpha_1} \cdots p_N^{\alpha_N})$ as $\sum_{k=1}^{N} \alpha_k u(p_k)$.

We can now state the theorem from which all the major resuts are derived.

ABEL'S THEOREM. *A necessary and sufficient condition for* ζ/ω *to be the divisor of a meromorphic function on* S *is that* $\deg \zeta = \deg \omega$ *and* $u(\zeta) \equiv u(\omega)$. $(\equiv)$ *meaning equal mod an integral linear combination of columns of the Riemann matrix.*

We shall use the Riemann-Roch theorem in the following form:

Let ζ be a divisor on S. Then $r[1/\zeta] = \deg \zeta - g + 1 + i(\zeta)$. Two immediate and important consequences being:

1) Let Δ be a divisor of an Abelian differential. Then the $\deg \Delta = 2g - 2$

2) There are exactly g orders n_i $(0 \leqq n_1 < n_2 \cdots < n_g \leqq 2g)$ that can be specified at each point $p \in S$ such that no non-constant meromorphic func-

tion exists on S whose only singularity is a pole of order n_i at p. These integers are called the gaps at p.

Given a general point $p \in S$ its gap sequence is $(1, 2, \cdots, g-1, g)$; however, there do exist points on S whose gap sequences omit some of these numbers. These points are called Weierstrass points. (It is well known that there exist only a finite number of Weierstrass points on S. In fact, if N be the number of such points counting multiplicities, then $2g+2 \leqq N \leqq g(g^2-1)$. The bounds given are due to Hurwitz [6] and the lower bound is assumed only in the hyperelliptic case.)

1. For the details of the construction of the Teichmueller (Torelli) spaces the reader is referred to the paper of Ahlfors [0] and Rauch [11, 12, 13, 14] and the reference there given. We remark that Rauch [13] showed that the Riemann matrix of a Torelli surface is independent of the particular representative chosen and (for the non-hyperelliptic surfaces) that a suitably chosen set of $3g-3$ of the periods give a complex analytic structure to the Torelli space. He then observed that these coordinates may be carried up to the Teichmueller space. Ahlfors [0] then showed that a different set of $3g-3$ quantities originally suggested by Schiffer and Spencer, provide a complex analytic structure for all of the Teichmueller space. We describe these quantities. Let $d\zeta$ be an Abelian differential of the first kind on S with simple zeros $p_0, p_1, \cdots, p_{2g-3}$. We coordinatize S with the g δ-periods of $d\zeta$ and the integrals of $d\zeta$ along suitably chosen paths from p_0 to the other $2g-3$ simple zeros. These coordinates will be denoted by $\mathfrak{J}_i$. For a more complete discussion of these quantities see Patt [10] and the references there given.

Patt [10] demonstrated the existence of a set of variations of a Teichmueller or Torelli surface depending on $3g-3$ complex parameters such that a neighborhood of the origin in the parameter space describes a complex analytic structure for a neighborhood of the original surface in the appropriate space. The above mentioned variations are performed in parameter discs about $3g-3$ points (suitably chosen) on the surface. We describe these variations.

Let $p_1, \cdots, p_{3g-3}$ be $3g-3$ suitably chosen points on the surface and let $D_1, \cdots, D_{3g-3}$ be the associated parameter discs. Furthermore, let $t_1, \cdots, t_{3g-3}$ be the local parameters in $D_1, \cdots, D_{3g-3}$. In each of the discs we consider the following variation of local parameter.

$$t_k^* = t_k + \frac{\epsilon_k}{2\pi i t_k} \quad k = 1, \cdots, 3g-3.$$

Utilizing the general formulae developed in his paper, Patt shows that

$$1)\quad \Delta\Pi_{ij} = \sum \epsilon_k \zeta_i'(t_k(p_k))\zeta_j'(t_k(p_k)) + 0(\epsilon^2) \quad i, j = 1, \cdots, g$$

$$2)\quad \Delta\mathcal{J}_I = \tfrac{1}{2\pi i}\sum \epsilon_k \eta_I'(t_k(p_k))\zeta'(t_k(p_k)) + 0(\epsilon^2) \quad I = g+1, \cdots, 3g-3$$

where Δ refers to the difference due to the variation, the prime refers to differentiation with respect to the local parameter, ζ is the integral of $d\zeta$ associated with the $\mathcal{J}_I$, and η_I is the normal integral of third kind with poles at p_0 and p_I, the p_I being the zeros of $d\zeta$.

These equations clearly yield that the π_{ij} and the $\mathcal{J}_I$ are analytic functions of the ϵ's and that

$$3)\quad \frac{\partial\Pi_{ij}}{\partial\epsilon_k} = \zeta_i'(t_k(p_k))\zeta_j'(t_k(p_k)),$$

$$4)\quad \frac{\partial\mathcal{J}_I}{\partial\epsilon_k} = \tfrac{1}{2\pi i}\eta_I'(t_k(p_k))\zeta'(t_k(p_k)).$$

Patt then defines his variation matrices and proves that for any n, n points on the surface can be chosen so as to maximize the rank of the variation matrix. An examination of the proof discloses that not only does there exist one set of n points which maximize, but rather the sets of n points which maximize are open in S^{3g-3}. The major result of Patt can now be stated as

LEMMA 1. *If $3g-3$ points are chosen on a Teichmueller or Torelli surface in such a way that the rank of the variation matrix is $3g-3$, then the ϵ's used in varying the surface at these points determine a complex analytic structure of dimension $3g-3$ in some sufficiently small neighborhood of the surface.*

THEOREM 1. *The product structure with S and $T^g(S)$ as factors may be endowed with a structure of complex analytic manifold in such a way that the space is an analytic fibre space over the base space $T^g(S)$ with fibre S_0 over $S_0(T) \in T^g(S)$.*

Proof. By virtue of the fact that the sets of $3g-3$ points which are referred to in Lemma 1 are open in S^{3g-3} it follows quite readily that corresponding to each point $P_0 \in S_0(T)$ there exists a neighborhood U of P_0 and a neighborhood N of $S_0(T)$ such that for $P \in U$ and $S(T) \in N$ the local coordinate at P is independent of the Teichmueller surface. Hence, each point $(P, S(T))$ of the product has such neighborhoods U and N and the transition functions are clearly analytic. The structure of the space is therefore seen to be generated by the local product structures in these neighborhoods.

The r-fold symmetric product S_r of a compact Riemann surface with itself is a complete non-singular manifold which may be identified with the Chow variety of integral divisors of degree r on S. Furthermore, if t_j is a local uniformizer of S at P_j $j=1,\cdots,r$, Andreotti [1] has shown that the elementary symmetric function of the t_j, the ϕ_j, serve as a set of local coordinates at $P_1\cdots P_r \in S_r$. These coordinates shall be called the Andreotti coordinates. Hence we may restate Theorem 1 as Theorem 1′.

The product structure with $S_r \times S_r$ and $T^g(S)$ as factors may be endowed with a structure of complex analytic manifold in such a way that the space is an analytic fibre space over the base space $T^g(S)$ with fibre $S_r' \times S_r'$ over $S'(T) \in T^g(S)$.

The space of Theorem 1′ is the space which shall play the crucial role in our further investigations. We shall denote this space by $W_{r,r}$.

2. We ask the reader to recall the definitions made concerning divisors in our introductory section. Note, that in general if ζ is of degree $r<g$, and *integral*, $r[1/\zeta]=1$. Therefore, by the Riemann-Roch theorem

$$1=r[1/\zeta]=\deg\zeta-g+1+i(\zeta) \text{ or } i(\zeta)=g-\deg\zeta=g-r.$$

We see therefore that a divisor of degree r has in general index of speciality equal to $g-r$.

Definition 1. A divisor ζ of degree r is termed special iff $i(\zeta)=g-r+1+m$ $m\geqq 0$.

The following lemma and its proof are due to Mayer.

LEMMA 2. *Let* $\omega\in S_r$ *and let* $u\colon S_r\to J(S)$. *Let* $\phi_1,\cdots,\phi_r$ *be Andreotti's coordinates at* $\omega\in S_r$ *and let* $\zeta_1,\cdots,\zeta_g$ *be local coordinates at* $x=u(\omega)\in J(S)$. *Then*

$$\operatorname{Rank}\left\|\frac{\partial\zeta_i}{\partial\phi_j}\right\|=g-i(\omega)\quad \begin{matrix} i=1,\cdots,g\\ j=1,\cdots,r\end{matrix}$$

where the rank is evaluated at the point $\omega\in S_r$.

Proof. We view the matrix as the matrix of a linear transformation on the space of Abelian differentials of the first kind to the space of r complex variables i.e.: $T(d\zeta_i)=\left(\frac{\partial\zeta_i}{\partial\phi_1},\cdots,\frac{\partial\zeta_i}{\partial\phi_r}\right)$ where ζ_i is the integral of $d\zeta_i$. The rank of the matrix is therefore nothing more than g minus the dimension of the null space. We therefore must show that the dimension of the null space is $i(\omega)$.

Case 1. Let $\omega \in S_r$ be such that $\omega = p_1, \cdots, p_r$, $p_i \neq p_j$ for $i \neq j$. Then $d\zeta \in \Omega(\omega)$ iff the divisor of $d\zeta$ is a multiple of ω which is true iff the zeros of $d\zeta$ include the points $p_1, \cdots, p_r$. If this is the case however $\frac{\partial \zeta}{\partial t_j} = 0$ at at ω for all j, $1 \leqq j \leqq r$, where t_j is the local uniformizer at p_j. This last remark however is just the statement of the fact that $d\zeta$ lies in null space of the linear transformation T' where $T'(d\zeta_i) = (\frac{\partial \zeta_i}{\partial t_1}, \cdots, \frac{\partial \zeta_i}{\partial t_r})$. Observe however that $T' = T \, \| \frac{\partial \phi_j}{\partial t_k} \|$. Since the p_j are all distant $\| \frac{\partial \phi_j}{\partial t_k} \|$ is non-singular. Indeed, its determinant is $\prod_{i<j} (t(p_i) - t(p_j))$. Hence we conclude that T and T' have the same null spaces.

Case 2. Let $\omega = p^r$ and let t be a local uniformizer at p. Near P, ζ may be expressed as a power series in t $a_0 + a_1 t + \cdots + a_n t^n + \cdots$. The induced function on C^r at the point $(p, \cdots, p)$ has the form

$$\zeta(t_1, \cdots, t_r) = r a_0 + a_1(t_1 + \cdots + t_r) + a_2(t_1^2 + \cdots + t_r^2) + \cdots.$$

Expanding the sums of powers in terms of the elementary symmetric functions and using the fact that $\phi_1, \cdots, \phi_r$ vanish at $\omega \in S_r$ we get that $\frac{\partial \zeta}{\partial \phi_j} = (-1)^{j+1} j a_j$ for $1 \leqq j \leqq r$. Therefore the $\frac{\partial \zeta}{\partial \phi_j} = 0$ for all j iff $a_j = 0$ for all j. Therefore $d\zeta$ has an r-th order zero at P iff $d\zeta$ is in the null space of T.

Observe, that we have now in particular porven the theorem for the case $r = 1$ and we proceed from here by induction. Assume the theorem proven for $r \leqq k - 1$ we show it's true for $r = k$. Let $\omega = p_1 \cdots p_k$. We can therefore write ω as $\omega' \omega''$ where $\omega' = p_1' \cdots p_{k'}'$, $\omega'' = p_1'' \cdots p_{k''}''$ and such that ω' and ω'' have no points in common. This decomposition into completely distinct components being permissable since if there does not exist $p_i \neq p_j$ for some i, j the previous case applies and the theorem is true. By the induction hypothesis the theorem is true for both ω' and ω''. Therefore $d\zeta \in \Omega(\omega')$ iff $\frac{\partial \zeta}{\partial \phi_j'} = 0$ and $d\zeta \in \Omega(\omega'')$ iff $\frac{\partial \zeta}{\partial \phi_k''} = 0$. Since ω' and ω'' are completely distinct and $\omega = \omega' \omega'' \Omega(\omega) = \Omega(\omega') \cap \Omega(\omega'')$. Hence we need only show that a necessary and sufficient condition for $\frac{\partial \zeta}{\partial \phi_i} = 0$ for all i is that $\frac{\partial \zeta}{\partial \phi_j'} = \frac{\partial \zeta}{\partial \phi_k''} = 0$ for all j and k. One half is trivial for

$$\frac{\partial\zeta}{\partial\phi_j'}=\Sigma\frac{\partial\zeta}{\partial\phi_i}\frac{\partial\phi_i}{\partial\phi_j'}\text{ and }\frac{\partial\zeta}{\partial\phi_k''}=\Sigma\frac{\partial\zeta}{\partial\phi_i}\frac{\partial\phi_i}{\partial\phi_k''}$$

Hence if $\frac{\partial\zeta}{\partial\phi_i}=0$ for all i then $\frac{\partial\zeta}{\partial\phi_j'}=\frac{\partial\zeta}{\partial\phi_k''}=0$. The converse follows by direct computation or by a geometric argument which the reader can find in Mayer's Princeton thesis [8].

LEMMA 3. *Let $\zeta,\omega\in S_r$ and let $u\colon S_r\to J(S)$. Let $\phi_1,\cdots,\phi_r$ be local coordinates at ζ and let $\psi_1,\cdots,\psi_r$ be local coordinates at ω. (Both the ϕ's and ψ's are understood to be Andreotti's coordinates). Then*

$$\operatorname{Rank}\left\|\frac{\partial\zeta_i}{\partial\phi_j}\,\frac{\partial\zeta_i}{\partial\psi_k}\right\|=g-i(\Delta)$$

where Δ is the least common multiple of ζ and ω. Hence

$$g-\min[i(\zeta),i(\omega)]\leqq\operatorname{Rank}\left\|\frac{\partial\zeta_i}{\partial\phi_j}\,\frac{\partial\zeta_i}{\partial\psi_k}\right\|\leqq g-i(\zeta\omega).$$

Proof. As in the proof of the previous lemma the rank of the matrix in question is just g minus the dimension of the null space of the appropriate linear transfromation. Hence $d\zeta$ is in the null space iff $\frac{\partial\zeta}{\partial\phi_j}=\frac{\partial\zeta}{\partial\psi_k}=0$. By the previous lemma however, this is true iff $d\zeta\in\Omega(\omega)$ and $d\zeta\in\Omega(\zeta)$. Hence the dimension of the null space is just $i(\Delta)$ where Δ is the least common multiple of ζ and ω. The inequalities on the rank now follow trivially from the fact that

$$i(\zeta\omega)\leqq i(\Delta)\leqq i(\zeta).$$

The remainder of this section is essentially concerned with proving one of our main theorems mentioned in the introduction.

THEOREM 2. *If ζ, ω are completely distinct equivalent divisors of degree r on S_0 then, considering the triple $(\zeta,\omega,S_0(T))$ as a point in $W_{r,r}$, for a sufficiently small neighborhood of $(\zeta,\omega,S_0(T))$, there is a g codimensional submanifold of $W_{r,r}$ (any point of which has projections on pairs of equivalent divisors on the surface which is the base point under the fibre) containing the point $(\zeta,\omega.S_0(T))$, which projects on a λ codimensional submanifold of the given neighborhood of $S_0(T)$. In addition $\lambda=i(\zeta\omega)$.*

We shall present the proof of this theorem in the form of a series of lemmas and then investigate some if its consequences.

LEMMA 4. *If from any* $g\times s$ *submatrix of rank* $s<g$ *one forms a* $g\times(s+1)$ *matrix by bordering with a column* $\zeta_1'(R),\cdots,\zeta_g'(R)$ *for a definite* $R\in S$ *then, an arbitrarily small shift of* R *will give the matrix rank* $s+1$.

Proof. Expand any $s+1$-st order determinant by its elements drawn from the bordering column. As a function of the point R the determinant is a non-identically vanishing Abelian differential of the first kind. Since an Abelian differential has only $2g-2$ zeros, it is clear that any neighborhood of a point must contain points which are not zeros of the differential. The differential under consideration is non-identically vanishing for by hypothesis at least one s-th order determinant was not zero.

LEMMA 5 (*Implicit Function Theorem*). *If the functions*

$$f_j(\omega_1,\cdots,\omega_k;z_1,\cdots,z_l) \qquad\qquad j=1,\cdots,k$$

are analytic functions of the $k+l$ *variables* $\omega_1,\cdots,\omega_k,z_1,\cdots,z_l$ *in a neighborhood of the origin and if* $f_j(0;0)=0$ $j=1,\cdots,k$ *and if*

$$\left|\frac{\partial(f_1,\cdots,f_k)}{\partial(\omega_1,\cdots,\omega_k)}\right|\neq 0 \text{ for } \omega=z=0$$

then, the equations $f_j=0$ *have a unique solution* $\omega_j=\omega_j(z_1,\cdots,z_l)$ vanishing *for* $z=0$ *and analytic in a neighborhood of the origin.*

Proof. Bochner and Martin [4].

Consider now a compact Riemann surface S and convenient Teichmueller surface $S(T)$ lying over S with the property that there exist on S completely distinct equivalent divisors of degree $r<g$. By Abel's theorem this is equivalent to the assertion that there exists on the surface a non-constant meromorphic function with r zeros and r poles. Letting ζ, ω denote the respective divisors, or equivalently the zeros and poles of the meromorphic function we have that $u(\zeta)\equiv u(\omega)$. This congruence relation is equivalent to the assertion that there exist integers m_k which satisfy the following equation:

$$u(\zeta)-u(\omega)-\sum_{k=1}^{2g}m_kR_k=0 \text{ where } R_k$$

denotes the k-th column of the Riemann matrix. Rewriting this equation $\zeta=p_1,\cdots,p_r$, $\omega=Q_1,\cdots,Q_r$ we obtain

$$\sum_{\alpha=1}^{r}u_i(p_\alpha)-u_i(Q_\alpha)-m_i-\sum_{j=1}^{g}n_j\Pi_{ij}=0 \qquad\qquad i=1,\cdots,g.$$

Letting $\phi_1, \cdots, \phi_r$ and $\psi_1, \cdots, \psi_r$ be Andreotti's coordinates at $\zeta, \omega \in S_r$ respectively we get

LEMMA 6.

$$W_i(\phi_1, \cdots, \phi_r, \psi_1, \cdots, \psi_r, \epsilon_1, \cdots, \epsilon_{3g-3}) = \sum_{\alpha=1}^{r} u_i(\phi_\alpha) - u_i(\psi_\alpha) - m_i - \sum_{j=1}^{g} n_j \Pi_{ij}$$

are analytic functions on $W_{r,r}$ where the ϵ's are Patt's coordinates and $u_i(\phi_\alpha)$, $u_i(\psi_\alpha)$ is a convenient functional notation.

Proof. Equations 3 and 4 of Section 1.

LEMMA 7. *The equations*

$$\sum_{\alpha=1}^{r} u_i(\phi_\alpha) - u_i(\psi_\alpha) - m_i - \sum_{j=1}^{g} n_j \Pi_{ij} = 0 \qquad i = 1, \cdots, g$$

define a g codimensional submanifold of $W_{r,r}$ containing the point $(\zeta, \omega, S(T))$ which projects onto a λ codimensional submanifold of a sufficiently small neighborhood of $S(T)$.

Proof. By Lemma 6 the g functions $W_i(\phi, \psi, \epsilon)$ are analytic functions of the $3g-3+2r$ variables and vanish at $\epsilon = \phi = \psi = 0$. Consider now the Jacobian matrix of the system of equations $W_i = 0$, $i = 1, \cdots, g$.

$$\left|\frac{\partial W_i}{\partial \phi_\alpha} \; \frac{\partial W_i}{\partial \psi_\alpha} \; \frac{\partial W_i}{\partial \epsilon_I}\right| \qquad \begin{array}{l} \alpha = 1, \cdots, r \\ I = 1, \cdots, 3g-3 \\ i = 1, \cdots, g \end{array}$$

From equations 3 and 4 of Section 1 we immediately deduce

$$\frac{\partial W_i}{\partial \epsilon_I} = \zeta_i'(t_I(R_I))\left[\tfrac{1}{2\pi i}\eta_{\zeta,\omega}'(t_I(R_I)) - \sum_{j=1}^{g} n_j \zeta_j'(t_I(R_I))\right].$$

Hence the Jacobian matrix now takes the form

$$\left|\frac{\partial W_i}{\partial \phi_\alpha} \; \frac{\partial W_i}{\partial \psi_\alpha} \; \zeta_i'(t_I(R_I))\left[\tfrac{1}{2\pi i}\eta_{\zeta,\omega}'\,(t_I(R_I)) - \sum_{j=1}^{g} n_j \zeta_j'(t_I(R_I))\right]\right|$$

Since ζ and ω are completely distinct we have by Lemma 2 that

$$\text{Rank} \left\|\frac{\partial W_i}{\partial \phi_\alpha} \; \frac{\partial W_i}{\partial \psi_\alpha}\right\| = g - i(\zeta\omega)$$

Let $i(\zeta\omega) = \lambda$ and consider now the right hand block of the Jacobian. We remark that the points R_I are chosen so as to satisfy the following three conditions:

a) No R_I is a zero of ζ_i'.

b) No R_I is a point of ζ or ω.

c) The R_I are chosen in such a way that the multipliers in the brackets are never zero.

That these demands may be made follows from the fact remarked earlier that the sets of $3g-3$ points which maximize the rank of the variation matrices are open in S^{3g-3}. Once again, using the fact that the set is open and Lemma 4 we may increase the rank of the left hand block by augmenting one column at a time from the right hand block until we obtain a maximal rank of g. Choosing now $g-\lambda$ linearly independent columns from the left hand block and augmenting λ columns from the right hand block, we obtain g coordinates which satisfy the Jacobian condition of the implicit function theorem. Invoking the implicit function theorem we can now solve our system of equations uniquely for these g coordinates with the property that the g coordinates vanish at the origin and furthermore, that these g coordinates are analytic functions of the remaining coordinates in a neighborhood of the origin. This, however, is the statement of the lemma. This concludes the proof of Theorem 2.

THEOREM 3. *The integer λ of Theorem 2 satisfies the following inequalities:*

$$\max(0, g-2r+1+m) \leqq \lambda \leqq g-r+1+m.$$

Therefore, for a sufficiently small neighborhood of $S_0(T)$ there is at most a λ codimensional submanifold of $T^g(S)$ on which there exist equivalent special divisors of degree r. The dimension of this maximal submanifold satisfies the inequalities

$$r+2g-4-m \leqq \dim \leqq V-2-m$$

where $V=2r+2g-2$ is the ramification index of the concrete r sheeted representation of S_0.

Proof. The inequalities involving λ are obtained in the obvious manner. Since $i(\zeta\omega) \leqq i(\zeta)$, and since ζ is assumed to be special $i(\zeta)=g-r+1+m$ we have $\lambda \leqq g-r+1+m$. The lower bound on λ comes from the Riemann-Roch count of losing at most one differential of the first kind with each additional point. The inequalities on the dimension follow immediately from the bounds on λ.

Consider now the general case of a special divisor with $m = 0$. We remark that it is sufficient to consider this case since $m > 0$ simply implies that there exists a special divisor of lower degree. It is clear that $V - 2 < 3g - 3$ iff $r < \frac{g+1}{2}$. Hence we see that if $r < \frac{g+1}{2}$ the special divisor is also special in the sense of moduli. In other words the sublocus of $T^g(S)$ which permit r sheeted representations $r < \frac{g+1}{2}$ is contained in a proper submanifold of $T^g(S)$. Furthermore, if $r > \frac{g+1}{2}$ we see that in *general* the divisor is not special in the sense of moduli. Hence we really need only examine the case of $r = \frac{g+1}{2}$ for odd genus.

THEOREM 4. *A special divisor of degree* $\frac{g+1}{2}$ *on* S, g *odd, is always special in the sense of moduli.*

Proof. Let ζ be a special divisor of degree $\frac{g+1}{2}$. Then ζ/ω is the divisor of a meromorphic function which projects S onto a $\frac{g+1}{2}$ sheeted covering of the z plane. By Abel's theorem $u(\zeta) \equiv u(\omega)$. The divisor ζ will therefore certainly be special in the sense of moduli iff $i(\zeta\omega) \geqq 1$. By hypothesis hower, $i(\zeta) = i(\omega)$ since $\zeta \sim \omega$, and $i(\zeta) \geqq \frac{g+1}{2}$ since ζ is special. Therefore let $\phi_1, \cdots, \phi_{\frac{g+1}{2}}$ and $\psi_1, \cdots, \psi_{\frac{g+1}{2}}$ be the respective linearly independent differentials of the first kind whose divisors are multiples of ζ and ω. Since there are more than a total of g differentials there must exist a linear relation between them say,

$$a_1\phi_1 + \cdots + a_{\frac{g+1}{2}}\phi_{\frac{g+1}{2}} \equiv b_1\psi_1 + \cdots + b_{\frac{g+1}{2}}\psi_{\frac{g+1}{2}}.$$

The left side vanishes at ζ hence so must the right side. However, the right side vanishes at ω. Therefore there exists at least one non-zero differential of the first kind which vanishes at $\zeta\omega$ and hence $i(\zeta\omega) \geqq 1$.

Combining the results of Theorem 4 and the remarks in the paragraph preceding it we have the following theorem:

THEOREM 5. *For even genus a divisor of degree less than* $\frac{g+2}{2}$ *is always* *special in the sense of moduli. For odd genus a divisor of degree less than* $\frac{g+3}{2}$ *is always special in the sense of moduli.*

Theorems 3, 4, and 5 seem to indicate that the property of permitting an r sheeted representation for $r \geqq \frac{g+2}{2}$ or $\frac{g+3}{2}$ depending on whether g is even or odd is not special in the sense of moduli. The proof would be complete if we were able to exhibit on S a $\frac{g+2}{2}$ or $\frac{g+3}{2}$ sheeted representation such that $i(\zeta\omega) = 0$ where ζ and ω are the zero and polar divisors of the function which projects S onto a $\frac{g+2}{2}$ or $\frac{g+3}{2}$ sheeted covering of the plane. In this connection the reader is referred to Meis [9] where examples are given.

We stress that our present remarks are not meant to imply that there cannot exist on a compact Riemann surface of genus g special divisors (in the sense of moduli) of degree greater than $\frac{g+1}{2}$, but rather that, in general, ones of higher degree are not special. As a matter of fact for $r > \frac{g+1}{2}$, a sufficient condition for $\zeta \in S_r$ to be special in the sense of moduli is that $i(\zeta) \geqq \frac{g+1}{2}$. The condition however is by no means necessary. Indeed, in the next section we shall show that there exist special divisors (in the sense of moduli) of degree $g-1$ and index 2.

We now give an example of the method and calculate the dimensions of some interesting subloci of $T^g(S)$.

We ask the reader to recall the definition of Weierstrass point given in the introduction. As we pointed out there, the Weierstrass sequence for a general point on S is the sequence of integers with the first g integers missing. We also indicated that there existed points on the surface whose Weierstrass sequences begin with integers less than $g+1$. These distinguished points were called Weierstrass points. Our present goal is to determine the dimension of the sublocus of $T^g(S)$ possessing Weierstrass points whose Weierstrass sequences begin with a fixed $r < g$.

LEMMA 8. *A necessary and sufficient condition that S possess Weierstrass points whose Weierstrass sequences begin with a fixed $r < g$ is that there exist on S a special divisor p^r such that $i(p^r) = g - r + 1$ and p must be such that $i(p^{r-1}) = i(p^r)$.*

Proof. By hypothesis there exists on S a non-constant meromorphic function with an r-th order pole at p and no other singularities. Hence by the Riemann-Roch theorem $r[\frac{1}{p^r}] = 2$. Since the Weierstrass sequence begins with r, $r[\frac{1}{p^{r-1}}] = 1$ therefore $r[\frac{1}{p^r}] - r[\frac{1}{p^{r-1}}] = 1$, or $i(p^{r-1}) = i(p^r)$.

THEOREM 6. *The sublocus of* $T^g(S)$ *possessing Weierstrass points whose sequences begin with a fixed* $r < g$ *is a complex analytic submanifold of* $T^g(S)$ *of dimension* $r + 2g - 3$ *in the case when* $i(Q_1 \cdots Q_r p) = i(Q_1 \cdots Q_r) - 1$ *and is contained in a subvariety of dimension* $r + 2g - 4$ *when* $i(Q_1 \cdots Q_r p) = i(Q_1 \cdots Q_r)$ *where* $u(Q_1 \cdots Q_r) \equiv u(p^r)$ *on* S.

Proof. By Lemma 8 the hypothesis of the theorem implies that there exists on any such surface a special divisor p^r whose index is $g - r + 1$. Therefore by the Riemann-Roch theorem there exists on the surface a non-constant meromorphic function with divisor $\frac{Q_1 \cdots Q_r}{p^r}$. By Abel's theorem we may write

$$u_i(p^r) = u_i(Q_1 \cdots Q_r) + \sum_{j=1}^{g} n_j \Pi_{ij} + m_i \qquad i = 1, \cdots, g.$$

Letting t be a local uniformizer at p and $\phi_1, \cdots, \phi_r$ be Andreotti's coordinates at $\phi_1 \cdots \phi_r \in S_r$, we may write

$$\sum_{\alpha=1}^{r} [r u_i(t) - u_i(Q_\alpha)] - \sum_{j=1}^{g} n_j \Pi_{ij} - m_i = 0 \qquad i = 1, \cdots, g$$

and view these equations as ranging over $W_{1,r}$. By the methods used in Theorems 2 and 3 we obtain a g codimensional submanifold of $W_{1,r}$ which projects onto a λ codimensional submanifold of a neighborhood of $S(T)$. $\lambda = g - r + 1$ or $g - r$ depending on whether $i(Q_1 \cdots Q_r) = i(Q_1 \cdots Q_r P)$ or not. In the latter case we obtain an $r + 2g - 3$ dimensional submanifold and in the former case an $r + 2g - 4$ dimensional subvariety of an $r + 2g - 3$ dimensional submanifold.

LEMMA 9. $i(Q_1 \cdots Q_r p) = i(Q_1 \cdots Q_r) - 1$ *iff* $r + 1$ *is a gap at* p. *Hence we may repalce the conditions of Theorem* 6 *with the condition that* $r + 1$ *be or not be a gap.*

Proof. Suppose $r + 1$ is a gap then $r[\frac{1}{p^r}] = r[\frac{1}{p^{r+1}}]$ and by the Riemann-Roch theorem $i(p^r) = 1 + i(p^{r+1})$. By hypothssis $u(p^r) \equiv u(Q_1 \cdots Q_r)$ therefore $u(p^{r+1}) \equiv u(Q_1 \cdots Q_r p)$. We pointed out in the introduction however that equivalent divisors have the same index. Hence $i(p^r) = 1 + i(p^{r+1})$ iff $i(Q_1 \cdots Q_r) = 1 + i(Q_1 \cdots Q_r p)$.

If we now in particular consider $r = 2$ we see that the sublocus of $T^g(S)$ possessing Weierstrass points whose sequences begin with 2 form a $2g-1$ dimensional submanifold in the case that 3 is a gap and a $2g-2$ dimensional sublocus of a $2g-1$ dimensional submanifold in the case that 3 is not a gap. However, the sublocus under consideration is the hyperelliptic sublocus and it is known that 3 is a grap for a hyperelliptic surface. Hence we have that the dimension of the hyperelliptic sublocus is indeed $2g-1$ confirming earlier knowledge. In this connection see Rauch [13].

Before concluding with this section we remark that the above examples are very illuminating in a certain respect. The estimates obtained on λ were the crudest possible, yet our examples indicate that the upper bound on λ can indeed be attained. The lower bound on λ can also obviously be attained by considering any general property. Hence we have that the bounds are best possible.

3. We now turn our attention to a particular first order theta function, the Reimann theta function $\theta(u-e) = \sum_N \exp(2\tilde{N}(u-e) + \tilde{N}TN)$ where

a) N runs over all integer vectors in C^g $n_1, \cdots, n_g$

b) $\tilde{N}$ is the transpose of N

c) T is a symmetric matrix with negative definite real part

d) u is a vector in C^g

e) e is a constant vector in C^g.

Clearly, given any S we can normalize its Riemann matrix so as to satisfy the conditions on T. Hence, we may associate with any compact Riemann surface a first order theta function. Furthermore, we may view the theta function as a multiplicative function on the surface via the map $u: S \to J(S)$. More explicitly, we consider the function $\theta(u(p) - e)\, p \in S$. Observe however that while the theta function was well defined on C^g we are not so fortunate in $J(S)$. That is $\theta(u_1)$ is not necessarily equal to $\theta(u_2)$ when $u_1 \equiv u_2$. Before proceeding we indicate the normal form of the Riemann matrix we shall be using:

$$\begin{Vmatrix} \pi i & 0 & \cdots & 0 & t_{11} & \cdots & t_{1g} \\ 0 & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & & \vdots & & \cdot \\ \cdot & & & \cdot\, 0 & \cdot & & \cdot \\ 0 & & \cdots & \pi i & t_{g1} & \cdots & t_{gg} \end{Vmatrix}$$

For general references about theta functions the reader is referred to Baker [2] or to the paper of Lewittes [7].

We now briefly summarize some of the well known facts concerning theta functions: While the theta function is not well defined on $J(S)$ its zeros are since if $u_1 \equiv u_2$ $\theta(u_1)$ differs from $\theta(u_2)$ by a non-zero exponential factor. Furthermore, as a multiplicative function on the Riemann surface $\theta(u(p)-e)$ is either identically zero or else has g zeros on the surface, $p_1, \cdots, p_g$, which satisfy the congruence relation $e \equiv u(p_1, \cdots, p_g) + K$. K is a vector of constants which depends only on the choice of base point for the integrals of first kind and on the homology basis. If $e \in C^g$ and if $\theta(e) \neq 0$ then $e \equiv u(\zeta) + K$ for a unique $\zeta \in S_g$ and $i(\zeta) = 0$. If $\theta(e) = 0$, let s be the smallest integer such that $\theta(W^s - W^s - e) \neq 0$, (where $W^s = u(S_s)$) then there is a $\zeta \in S_{g-1}$, $i(\zeta) = s$, such that $e \equiv u(\zeta) + K$. Also, all partials derivatives of θ of order less than s vanish at e while at least one partial of order s does not vanish at e. In regard to the vector K of constants just mentioned we remark that a divisor Δ of degree $2g-2$ is the divisor of an Abelian differential of first kind iff $u(\Delta) \equiv -2K$. More generally we may consider a first order theta function with characteristics. Let G, H be two vectors in C^g. We define $\theta[\begin{smallmatrix} G \\ H \end{smallmatrix}](u) = \exp(\tilde{G}TG + 2\tilde{G}u + 2\pi i\tilde{G}H)\theta(u + TG + \pi iH)$. As a function on the Riemann surface $\theta[\begin{smallmatrix} G \\ H \end{smallmatrix}](u)$ is either identically zero or has g zeros on the surface which satisfy the congruence relation.

$$u(p_1, \cdots, p_g) \equiv -\pi iH + T(1-G) - K.$$

We now consider those points $e \in C^g$ which have the property $2e \equiv 0$. Clearly, modulo the group of integral linear translations defined by the columns of the Riemann matrix there are 2^{2g} such points. This remark follows from the fact that in order for $2e \equiv 0$, e must be of the form $T\frac{\epsilon}{2} + \pi i\frac{\epsilon'}{2}$ where T is the right hand block of the normal Riemann matrix and ϵ, ϵ' and are integer vectors in C^g. Moduo the group considered, in order to obtain all points with the desired property it is sufficient to let the entries of ϵ and ϵ' be zero or one. ince there are 2^{2g} ways of doing this there are 2^{2g} such points. Points with the property that $2e \equiv 0$ will be termed half periods.

We have seen that each half period determines two integer vectors in C^g, ϵ, ϵ' with zero and one entries. The half period is called even or odd depending on whether the scalar product $\epsilon \cdot \epsilon'$ is even or odd. The reason for the even-odd terminology is the following: Let $\epsilon/2$, $\epsilon'/2$ be the characteristics of a theta function. The function is then even or odd depending on whether $\epsilon\cdot\epsilon'\equiv 0$ or $1 \bmod 2$ [5].

It follows immediately from the definition that

$$\theta\begin{bmatrix}\frac{\epsilon}{2}\\ \frac{\epsilon'}{2}\end{bmatrix}(0)=\exp\left(\frac{\bar{\epsilon}}{2}T\frac{\epsilon}{2}+\frac{\pi i}{2}\bar{\epsilon}\epsilon'\right)\theta(e)$$

where

$$e=T\frac{\epsilon}{2}+\pi i\frac{\epsilon'}{2}.$$

Therefore a theta function with half period characteristics vanishes at the origin iff the theta with zero characteristic vanishes at the half period. An odd function vanishes at the origin. Hence the theta function always vanishes at an odd half period. An even function does not generally vanish at the origin. Hence we may inquire as before, whether the special property of an even theta vanishing at the origin is special in the sense of moduli.

THEOREM 7. *Let θ be the sublocus of $T^g(S)$ satisfying $\theta(e)=0$ e an even half period. Then θ is contained in at most a $3g-4$ submanifold of $T^g(S)$. Hence the theta function vanishing at an even half period is special in the sense of moduli.*

Proof. It is a consequence of the properties of the theta function that the vanishing of the theta function at an even half period is equivalent to the existence on the surface of equivalent special divisors ζ, ω of degree $g-1$. Furthermore $i(\zeta)=i(\omega)=2$ and $i(\zeta\omega)=1$. As usual let $\phi_1,\cdots,\phi_{g-1}$ and $\psi_1,\cdots,\psi_{g-1}$ be the Andreotti coordinates at ζ and ω respectively and view the equations of Abel's theorem as ranging over $W_{g-1,g-1}$.

We now use the estimates obtained in Theorem 3 to obtain $1\leqq\lambda\leqq 2$ and this gives us the theorem. We remark that λ is one or two depending on the index of the least common multiple of ζ and ω. It is not at all clear that ζ and ω are completely distinct.

THEOREM 8. *Completely distinct equivalent special divisors of degree $g-1$ are special in the sense of moduli iff they arise from a theta function vanishing at a half period.*

Proof. For $\zeta,\omega\in S_{g-1}$, $\zeta\sim\omega$, $i(\zeta\omega)=0$ or 1 depending on whether $\zeta\omega$ is not or is the divisor of a differential of the first kind on S. In the former case $\lambda=0$ and hence the divisor is not special in the sense of moduli. In the latter case $\lambda = 1$ and hence the divisor is special in the sense of moduli. By the argument of Theorem 8 however we get $e \equiv u(\zeta) + K$, e a half period and hence $\theta(e) = 0$.

We now pose the following question: Suppose that not only does the theta function vanish at an even half period but so do its first $2k-1$ partial derivatives. One would like to say that the sublocus of $T^g(S)$ satisfying this condition is contained in a $2k-1$ co-dimensional submanifold of $T^g(S)$; however, life is not so pleasant. What one may say is embodied in the following theorem:

THEOREM 9. *Let θ_{2k-1} be the sublocus of $T^g(S)$ satisfying $\theta(e) = 0$ and the first $2k-1$ partials of θ at $e = 0$ for e an even half period. Then θ_{2k-1} is contained in a λ co-dimensional submanifold of $T^g(S)$, ($\lambda = i(\Delta)$, the least common multiple of ζ and ω) where $1 \leqq \lambda \leqq 2k$.*

Proof. Same as Theorem 7 replacing 2 by $2k$ for $i(\zeta)$ and $i(\omega)$.

We remark that one cannot hope for better answers using this method. The reason being that simply writing down the equations of Abel's theorem $u(\zeta) \equiv u(\omega)$ admits a wider class of surfaces than one wants. For example, if the theta function and its first 3 partials vanish at an even half period on a surface of genus 7, simply writing down the equations $u(\zeta) \equiv u(\omega)$, $i(\zeta) = 4$ does not exclude from the sublocus defined, a surface on which the theta function vanishes at an odd half period and $i(\zeta) = 3$.

We conclude this paper with the proof of the following theorem. "A necessary and sufficient condition for a compact Riemann surface of genus 4 to be hyperelliptic is that the associated theta function vanish at two even half periods."

The statement of this theorem appears as an exercise in Baker [2]; however, no complete proof has yet appeared in the literature. We present here a proof which is essentially a patching up of the proof offered by Weber [18].

As we have previously indicated, the vanishing of the theta function at an even half period is equivalent to the existence on the surface of special divisors ζ, ω $\deg \zeta = \deg \omega = g-1$ such that ζ^2, $\zeta\omega$, ω^2 are divisors of Abelian differentials of first kind and such that $u(\zeta) \equiv u(\omega)$.

LEMMA 10. *Let the theta function vanish at an even half period and let ϕ_1, ϕ_2, ϕ_3 be three Abelian differentials of first kind whose divisors are ζ^2, $\zeta\omega$, ω^2. Then ϕ_1, ϕ_2, ϕ_3 are linearly independent. In particular, therefore, the theta function vanishing at two even half periods implies that there exists* *on the surface two partial bases for the differentials of first kind* ϕ_1, ϕ_2, ϕ_3; ψ_1, ψ_2, ψ_3 *with the relations* $\phi_2{}^2 = \phi_1\phi_3$; $\psi_2{}^2 = \psi_1\psi_3$.

Proof. Since $\omega \neq \zeta$ ϕ_3 has a zero which ϕ_1 does not and conversely. Furthermore, the zeros of ϕ_1 and ϕ_3 are also zeros of ϕ_2. Assume now $a_1\phi_1 + a_2\phi_2 + a_3\phi_3 \equiv 0$. Let p_1 be a zero of ϕ_1 which is not a zero of ϕ_3 hence $a_3 = 0$. Similarly $a_1 = a_2 = 0$. The same argument applies to ψ_1, ψ_2 and ψ_3.

LEMMA 11. *We assume the following for S genus* 4.

a) *ϕ_1 has three distinct zeros*

b) *ψ_2 has one common zero with ϕ_1 and at most two common zero with ϕ_1.*

c) *ϕ_3 and ϕ_1 have no common zeros.*

Proof. ϕ_1 will have 3 distinct zeros iff the divisor ζ of the previous lemma consists of 3 distinct points. Consider the function f whose divisor is ξ_1/ω_1. Furthermore let f_* be the family of functions $f - f(p^*)p^* \notin \omega_1$. Let $\xi_1{}^*$ be the divisor of zeros of f_*. Then clearly by Abel's theorem $u(\xi_1{}^*) \equiv u(\omega_1) \equiv u(\xi_1)$. Since f has only a finite number of multiple points $\xi_1{}^*$ consists of 3 distinct points for some $P^* \in S$. A similar argument gives the first half of b. The second half of b, and c may be assmed, for otherwise, by the sufficiency of Abel's theorem S is hyperelliptic.

LEMMA 12. *Under the assumptions of Lemma 4 the differentials ϕ_1, ϕ_2, ϕ_3, Ψ_2 are linearly independent.*

Proof. Assume $a_1\phi_1 + a_1\phi_2 + a_3\phi_3 + a_4\Psi_2 \equiv 0$. Let $\phi_1(P_1) = \Psi_2(P_1) = 0$. Then $\phi_2(P_1) = 0$ and hence $a_3 = 0$. Let P_2 be such that $\phi_1(P_2) = 0$, $\Psi_2(P_2) \neq 0$. Then $\phi_2(P_2) = 0$ and hence $a_4 = 0$. By Lemma 3, ϕ_1, ϕ_2 are independent hence $a_1 = a_2 = 0$.

LEMMA 13. *The differential Ψ_2 may be chosen in such a way that it has no zeros in common with ϕ_2 and remains linearly independent from ϕ_1, ϕ_2, ϕ_3.*

Proof. Utilizing the arugment used in proof of a) in Lemma 4 we can initially arrange things in such a way that Ψ_2 and ϕ_2 have no zeros in common. As per b) of Lemma 4 we can then arrange that Ψ_2 have only one zero in common with ϕ_1. Call the new differential $\Psi_2{}^*$. The divisor of $\Psi_2{}^*$ is $\xi_2\omega_2{}'$ where $\omega_2{}'$ has one point in common with ξ_1. The general Wronskian argument will now give that $\Psi_2{}^{*\prime}$ with divisor $\xi_2\omega_2{}''$ with $\omega_2{}''$ in a sufficiently small neighborhood of $\omega_2{}'$ (and no zeros in common with ϕ_2) has the same linear properties as $\Psi_2{}^*$. Hence we have $\Psi_2{}^{*\prime}$ is linearly independent of ϕ_1, ϕ_2, ϕ_3.

THEOREM 10. *Let S be a compact Riemann surface of genus 4. Then the theta function vanishing at 2 even half periods implies that there exists on the surface a basis for the differentials of the first kind ϕ_1, ϕ_2, ϕ_3, Ψ_2 with the relations*

$$3) \qquad \phi_2^2 = \phi_1\phi_3 \qquad \Psi_2^2 = \Psi_1\Psi_3$$

where Ψ_1, Ψ_2, Ψ_3 are linearly independent and Ψ_2 has no common zeros with ϕ_2.

THEOREM 11. *Let S be a compact Riemann surface of genus 4. Then S is hyperelliptic iff the theta function associated with S vanishes at two even half periods.*

Proof. Let $P_1P_2P_3$ and $Q_1Q_2Q_3$ be the zeros of ϕ_1 and ϕ_3 respsctively then consider once again equation 3 which we write as

$$\psi_2^2 \equiv (a_1\phi_1 + a_2\phi_2 + a_3\phi_3 + a_4\psi_2)(b_1\phi_1 + b_2\phi_2 + b_3\phi_3 + b_4\psi_2).$$

At the points P_1, P_2, P_3; Q_1, Q_2, Q_3, respectively the identity reduces to the respective equations

$$4) \qquad (1 - a_4b_4)\Psi_2^2 - (a_3b_4 + b_3a_4)\phi_3\Psi_2 - a_3b_3\phi_3^2 = 0$$

$$5) \qquad (1 - a_4b_4)\Psi_2^2 - (a_1b_4 + a_4b_1)\phi_1\Psi_2 - a_1b_1\phi_1^2 = 0$$

which must be satisfied at the points P_1, P_2, P_3; Q_1, Q_2, Q_3 respectively. Both these equations cannot vanish identically for if both did we would have $a_1 = b_1 = a_3 = b_3 = 0$ which will contradict the linear independence of Ψ_1, Ψ_3, Ψ_2. Hence at least one of these equations does not vanish identically. Assume that it is equation 4 which does not vanish identically. It is sufficient to examine two cases.

a) $\quad (1 - a_4b_4) = 0 \qquad$ b) $\quad (1 - a_4b_4) \neq 0$.

If a) we have $\phi_3[(a_3b_4 + a_4b_3)\Psi_2 + a_3b_3\phi_3] = 0$ at $P_1P_2P_3$. Since $\phi_3 \neq 0$ at any of the points we have $a_3b_3\phi_3 + (a_3b_4 + a_4b_3)\Psi_2$ vanishes at $P_1P_2P_3$. But then $i(P_1P_2P_3) = 3$ contradicting the fact that $i(P_1P_2P_3) = 2$. Hence we may assume $1 - a_4b_4 \neq 0$. Equation 4 then can be factored as follows

$$\left(\Psi_2 - \phi_3\left[\frac{-\lambda_2 + \sqrt{\lambda_2^2 - 4\lambda_1\lambda_3}}{2\lambda_1}\right]\right)\left(\Psi_2 - \phi_3\left[\frac{-\lambda_2 - \sqrt{\lambda_2^2 - 4\lambda_1\lambda_3}}{2\lambda_1}\right]\right)$$

and this must vanish at $P_1P_2P_3$, where λ_1, λ_2, λ_3 are the coefficients of Ψ_2^2, $\phi_3\Psi_2$, ϕ_3^2 respectively. One of the factors must therefore vanish at two of the points, P_1P_2. But then $i(P_1P_2) = 3$ since both factors are independent of ϕ_1 and ϕ_2. The proof of the sufficiency will be complete if we observe that $i(P_1P_2) = 3$ implies that there exists on S a non-constant function f with poles at P_1 and P_2. That this is true follows from the Riemann-Roch theorem. The necessity follows from the well known fact that on a hyperelliptic surface of genus 4 the theta function vanishes at 10 even half periods.

BELFER GRADUATE SCHOOL OF SCIENCE,
YESHIVA UNIVERSITY.

REFERENCES.

[0] L. Ahlfors, "The complex analytic structure of the space of closed Riemann surfaces," *Analytic Functions*, Princeton University Press, 1960, pp. 45-66.

[1] A. Andreotti, "On a theorem of Torelli," *American Journal of Mathematics*, vol. 80 (1958), pp. 801-828.

[2] H. F. Baker, *Abelian Functions*, Cambridge University Press (1897).

[3] L. Bers, "Spaces of Riemann surfaces," *Proceedings of the International Congress of Mathematics*, 1958, Cambridge.

[4] Bochner and Martin, *Several Complex Variables*, Princeton, 1948.

[5] F. Conforto, *Abelsche Funktionen und Algebraische Geometrie*, Springer Verlag, 1956, pp. 103-104.

[6] A. Hurwitz, "Uber Algebraische Gebilde mit Eindeutigen Transformationem in sich," *Mathematische Werke*, Bd. I, Birkhauser Basll, 1932, pp. 392-430.

[7] J. Lewittes, "Riemann surfaces and the theta function," *Acta Mathematica*, vol. 111 (1964), pp. 37-61.

[8] A, Mayer, *On the Jacobi Inversion Theorem*, Princeton thesis, 1961.

[9] Th. Meis, "Die minimale Blätterzahl der Konkretisierung einer kompakten Riemannschen Flache," *Schriftenreihe des Mathematisches Institut der Universität*, Munster, 1960.

[10] C. Patt, "Variations of Teichmueller and Torelli surfaces," *Journal d'Analyse Math.*, vol. 11 (1963), pp. 221-247.

[11] H. E. Rauch, "On the transcendental moduli of algebraic Riemann surfaces," *Proceedings of the National Academy of Science, U. S. A.*, vol. 41 (1955), pp. 42-49.

[12] ———, "On the moduli of Riemann surfaces," *Proceedings of the National Academy of Science, U. S. A.*, vol. 41 (1955), pp. 236-238.

[13] ———, "Weierstrass points, branch points, and the moduli of Riemann surfaces," *Communications on Pure and Applied Mathematics*, vol. 12 (1959), pp. pp. 543-560.

[14] ———, "A transcendental view of the space of algebraic Riemann surfaces," *Bulletin of the American Mathematical Society*, vol. 71, no. 1 (Jan. 1965), pp. 1-39.

[15] B. Riemann, "Theorie der Abelschen Funktionen," *Gesammelte Mathematische Werke*, Zweite Auflage, Dover, N. Y.

[16] G. Springer, *Introduction to Riemann Surfaces*, Addison Wesly, Reading, 1957.

[17] O. Teichmueller, "Extremale Quasikonformal Abbildungen und Quadratische Differentiale," *Abhandlungen Preuss, Akademie von Wissenschafter Math-Naturwiss*, Kl. 22, 1940.

[18] H. Weber, *Uber Gewisse in der Theorie der Abelschen Funktionen*, Auftretende Ausnahmfälle, Konigsberg, 1877.

DECOMPOSITION THEOREM FOR BOUNDED UNIFORMLY CONTINUOUS FUNCTIONS ON A GROUP.

By A. W. KNAPP.[1]

Introduction. The main theorem of this paper states roughly that the Banach algebra of all bounded left-uniformly continuous functions on a topological group is the direct sum of a space of functions which in a sense vanish at infinity and a space of functions which "oscillate regularly." The first space is any proper self-adjoint closed ideal I maximal with respect to the property of being invariant under left translation. (Such ideals exist by Zorn's Lemma.) The second space is a self-adjoint closed algebra A containing the constants and invariant under left translation.

The decomposition is not canonical, even if I is held fixed. But any A of the type described above that is a direct summand with some such I is a direct summand with every such I, and vice-versa. Each such algebra A contains all almost periodic functions and even the wider class of right distal functions, which are defined in [5].

The set-theoretic union of all the algebras A turns out to be the set of right minimal functions, which are defined intrinsically in Section 3 in terms of limits of their translates. Also in Section 3, right minimal functions are characterized in terms of the ϵ-translation numbers familiar from the theory of almost periodic functions. It is this characterization which gives the sense in which right minimal functions "oscillate regularly."

For studying minimal functions, we shall use shift operators, which are developed in Section 2 and which essentially are just limits of translation operators. This device was suggested partly by the work of Bochner [2]. Shift operators will be used extensively at a later date when we elaborate on the results announced in [5].

I am indebted to Robert Strichartz for many helpful conversations in connection with this paper.

1. Decomposition theorem.

Let G be a topological group with identity

Received September 1, 1965.

[1] This research was sponsored by a fellowship from the National Science Foundation, and the preparation of this paper was supported by Air Force Office of Scientific Research Grant AF-AFOSR 335-63.

e. If f is a complex-valued function on G, then $\| f \|$ denotes the supremum norm of f and f_g is the left translate of f with $f_g(h) = f(gh)$. Such a function is left-uniformly continuous if $\lim_{g \to e} \| f_g - f \| = 0$. The set of bounded left uniformly continuous functions is denoted LUC; LUC is a self-adjoint (i. e., closed under conjugation) Banach algebra containing the constants and closed under left and right translation.

The theorem to follow is the decomposition theorem. Its proof can be simplified when A_0 is the algebra of constants with the aid of the machinery developed by Ellis in [3] and [4]. The price is the loss of both the elementary nature of the proof and the motivation for considering shift operators.

THEOREM 1-1. *Let I be a proper self-adjoint ideal in* LUC *which is maximal with respect to the property of being invariant under left translation. (Such ideals exist by Zorn's Lemma and are necessarily closed.) Let A_0 be any algebra in* LUC *which*

(a) *contains* 1
(b) *is invariant under left translation and conjugation*
(c) *satisfies $A_0 \cap I = 0$*
(d) *is such that the projection from $A_0 \oplus I$ onto A_0 preserves order (i. e., $a + i \geqq 0$ implies $a \geqq 0$).*

Condition (d) *on A_0 is automatically satisfied if A_0 is closed under uniform limits and satisfies* (a), (b), *and* (c). *Then there exists an algebra A containing A_0 and closed under left translation, conjugation, and uniform limits such that* LUC $= I \oplus A$.

Proof. (1) We first prove that if A_0 is uniformly closed and satisfies (a), (b), and (c), then A_0 satisfies (d). Thus let A_0 be closed and let $a + i \geqq 0$ with $a \in A$ and $i \in I$. Then $\operatorname{Re} a \in A$ and $\operatorname{Re} i \in I$, and

$$a + i = \operatorname{Re} a + \operatorname{Re} i.$$

Since $A_0 \cap I = 0$, a and i must be real. Since A_0 is self-adjoint and uniformly closed, A_0 is closed under absolute value and hence maxima. Since $0 \in A_0$, $\max(a, 0) \in A_0$. If we can show that $\min(a + i, i) \in I$, then the equality

$$a + i = \max(a, 0) + \min(a + i, i)$$

together with the fact that $A_0 \cap I = 0$ shows that $a = \max(a, 0)$ or $a \geqq 0$. Now I can be realized as the set of continuous functions on the maximal ideal space of LUC which vanish on some closed set. On that set, $i = 0$ and $a + i \geqq 0$. Hence on that set, $\min(a + i, i) = 0$, and thus $\min(a + i, i) \in I$.

(2) We now produce A. Partially order by inclusion the set of all algebras in LUC which contain A_0 and which satisfy (a), (b), (c), and (d). Any chain in the set clearly has its union as an upper bound in the set. By Zorn's Lemma take A to be a maximal element.

(3) Next we construct the map T that will be the projection onto A. The sum $A \oplus I$ is a self-adjoint normed algebra which has an identity and is invariant under left translation. Define ρ_0 on $A \oplus I$ to be evaluation of the A part at e. Then ρ_0 is a multiplicative linear functional on $A \oplus I$ which commutes with conjugation. Property (d) implies that ρ_0 is bounded by 1. By uniform continuity extend ρ_0 to ρ_1 defined on the closure of $A \oplus I$ in LUC. Then choose ρ to be a multiplicative linear functional commuting with conjugation which extends ρ_1 to all of LUC; we have $\| \rho \| = 1$. Define T by

$$Tf(g) = \rho(f_g) \quad \text{for } g \in G.$$

Then T is a homomorphism which commutes with left translation and conjugation. As a map into the space of bounded functions on G, its norm is one. Since

$$\| (Tf)_g - Tf \| = \| Tf_g - Tf \| \leqq \| f_g - f \|,$$

its range is in LUC. Let B be its image in LUC. Then B is a self-adjoint algebra with identity and is invariant under left translation. Therefore B satisfies (a) and (b).

(4) We shall show $\ker T = I$. If $f \in I$, then $f_g \in I$ for every g and hence

$$Tf(g) = \rho(f_g) = \rho_0(f_g) = 0.$$

Thus $\ker T \supseteq I$. In the reverse direction, $\ker T$ is a left-invariant self-adjoint ideal in LUC. By the maximality of I, $\ker T = \text{LUC}$ or $\ker T = I$. The first choice is impossible since $T1 = 1$.

(5) Note that $B \supseteq A$ and that T is the identity on A. In fact, if $f \in A$, then

$$Tf(g) = \rho(f_g) = \rho_0(f_g) = f_g(e) = f(g),$$

so that $Tf = f$, $f \in B$, and T is the identity on A.

(6) We claim that $B \cap I = 0$. In fact, $T^{-1}(B \cap I)$ is a self-adjoint ideal in LUC invariant under left translation, and it contains I by (4). By the maximality of I, $T^{-1}(B \cap I) = \text{LUC}$ or $T^{-1}(B \cap I) = I$. If $T^{-1}(B \cap I) = \text{LUC}$, then

$$B = T(\mathrm{LUC}) = TT^{-1}(B \cap I) = (B \cap I) \cap B = B \cap I$$

and $1 \in I$, a contradiction. Thus $T^{-1}(B \cap I) = I$ and

$$0 = T(I) = T(T^{-1}(B \cap I)) = (B \cap I) \cap B = B \cap I.$$

(7) Next we show that T induces an isometry of LUC/I with B and hence B is closed in LUC. We are to show that

$$\inf_{i \in I} \| f + i \| = \| Tf \|.$$

In one direction

$$\| Tf \| = \| T(f+i) \| \leqq \| T \| \, \| f+i \| = \| f+i \|,$$

and hence $\| Tf \| \leqq \inf_i \| f+i \|$. In the other direction we shall produce an $i \in I$ such that $\| f+i \| = \| Tf \|$. If $\rho_g(f) = \rho(f_g)$, we have

$$\| Tf \| = \sup_{g \in G} | \rho_g(f) |.$$

Let C be the closure in the maximal ideal space of LUC of $\{\rho_g\}$. Then

$$\| Tf \| = \sup_{x \in C} | x(f) |.$$

The functions vanishing on C form a closed self-adjoint left-invariant ideal, and every function in I is annihilated by every ρ_g. By the maximality of I, I is the set of functions in LUC which vanish on C. Extend the restriction of f to C to a continuous function g defined on the entire maximal ideal space of LUC without an increase in norm, and put $i = g - f$. Then $i \in I$ and

$$\| f+i \| = \| g \| = \| f \mid C \| = \sup_{x \in C} | x(f) | = \| Tf \|.$$

Hence B is uniformly closed.

(8) Since B is closed and since B satisfies (a), (b), and (c), step (1) shows that B satisfies (d). By the maximality of A, $B = A$ and A is therefore closed. By (5), T is the identity on A, and A is the image of T since $B = A$. Therefore T is a projection. Its image is A and its kernel, by (4), is I. Thus $\mathrm{LUC} = I \oplus A$, and the proof is complete.

The corollary below was stated and proved by R. Strichartz. It shows that any A that is a direct summand with some I is a direct summand with every I and vice-versa.

COROLLARY 1-2. *If* $\mathrm{LUC} = I \oplus A$ *and* $\mathrm{LUC} = I' \oplus A'$ *are two decompositions given by Theorem* 1-1, *then also* $\mathrm{LUC} = I' \oplus A$.

Proof. Since A and LUC/I are isomorphic, A contains no proper non-zero self-adjoint ideals invariant under left translation. Thus $I' \cap A = 0$ or $I' \cap A = A$. And since 1 is not in I', we must have $I' \cap A = 0$. Therefore by Theorem 1-1 there is a closed A_1 with $A_1 \supseteq A$ and $\mathrm{LUC} = I' \oplus A_1$. Reversing the argument, we can find a closed A_2 with $A_2 \supseteq A_1$ and $\mathrm{LUC} = I \oplus A_2$. Since $A_2 \supseteq A$ and $\mathrm{LUC} = I \oplus A_2 = I \oplus A$, we have $A_2 = A$. Then $A_1 = A$ and $\mathrm{LUC} = I' \oplus A$.

2. **Shift operators.**

The way the projection homomorphism T of LUC onto A in Theorem 1-1 was constructed was to take a certain multiplicative linear functional, translate it around the group, and fit the pieces together. Now any multiplicative linear functional is a weak-star limit of evaluations, and it follows that from any such functional this construction always leads to homomorphisms of LUC which are limits of right-translation operators in an appropriate sense. In this section we introduce the set of all such limiting operators.

We consider nets $\{g_n\}$ in G (i. e., functions on a directed set) such that $\lim_n f(gg_n)$ exists pointwise for every g in G and for every f in LUC. Every convergent net, for instance, has this property. Call two such nets $\{g_n\}$ and $\{h_m\}$ equivalent if

$$\lim_n f(gg_n) = \lim_m f(gh_m)$$

pointwise for every g in G and f in LUC. The operator that sends f into its limiting value for that equivalence class is called the *shift operator* associated to the class. Shift operators are denoted T_α, T_β, etc. Every *right*-translation operator is a shift operator.

Each shift operator is a homomorphism of norm one on LUC, each commutes with conjugation and *left* translation, and each has range in LUC. Since an iterated limit of nets can always be realized as a single limit, the composition of two shift operators is again a shift operator.

The class of shift operators is a set because distinct shift operators are distinct homomorphisms of LUC. A subbase for a topology on this set consists of all sets of the form

$$N(T_\alpha, f, x, \epsilon) = \{T_\beta \mid |T_\beta f(x) - T_\alpha f(x)| < \epsilon\}.$$

In this topology a net $\{T_{\alpha_n}\}$ converges to T_α if, for every f in LUC, the net of functions $\{T_{\alpha_n} f\}$ converges pointwise to $T_\alpha f$.

This topology is the one induced by the multiplicative linear functionals under the construction described in the first paragraph of this section. There is a one-one correspondence between the multiplicative linear functionals and the shift operators defined as follows: If a functional ρ is given, put $T_\alpha f(g) = \rho(f_g)$; any net of evaluations that converges weak-star to ρ is a net in G that defines T_α. On the other hand, if a shift operator T_α is given, put $\rho(f) = T_\alpha f(e)$. We thus get a one-one onto correspondence, and the topology on the set of shift operators is the topology that makes the correspondence a homeomorphism.

Since the weak-star topology on the set of multiplicative linear functionals is compact Hausdorff, so is the topology on the set of shift operators. Moreover, if $\lim T_{\alpha_n} = T_\alpha$, then clearly $\lim T_{\alpha_n} T_\beta = T_\alpha T_\beta$ for fixed T_β. That is, for fixed T_β the map $T_\alpha \to T_\alpha T_\beta$ is continuous. We have proved

THEOREM 2-1. *The set of shift operators is a semigroup under composition and is a compact Hausdorff space. For fixed T_β the map $T_\alpha \to T_\alpha T_\beta$ is continuous.*

Multiplication in the semigroup is not jointly continuous unless every function in LUC is almost periodic. The compactness of the semigroup implies that any net in G has a subnet which defines a shift operator.

3. Minimal functions. We shall be concerned in this section with the set of all algebras A which can occur in Theorem 1-1 if A_0 is taken as the algebra of constants. Evidently any such A is isomorphic to LUC/I, and consequently A contains no proper non-zero self-adjoint left-invariant ideals. This fact has an implication for the maximal ideal space of A, and we begin at this point.

We first recall the definition of a flow and the connection between functions and flows. A *flow* (X, G) for G is a jointly continuous group action of G on a compact Hausdorff space. (We shall suppose the group acts on the left.) If p is a point of X, then for each continuous complex-valued function f on X, the function $\bar{f}$ defined on G as

$$\bar{f}(g) = f(gp)$$

is in LUC, and the set of all such functions $\bar{f}$ for fixed p forms a self-adjoint Banach algebra containing the constants, invariant under left translation, and having maximal ideal space the closure of the orbit of p in X. We shall say $\bar{f}$ arises from the orbit of p. Conversely if such an algebra is given, then the left action of G on itself extends in a nautral way to a jointly continuous action of G on the maximal ideal space of the algebra. These two constructions are inverse to one another if p is chosen to be the evaulation-at-e functional in the maximal space.

We say that $(Y, G) = \pi(X, G)$ is a *subflow* of (X, G) if (X, G) and (Y, G) are flows and if π is a continuous map of X onto Y commuting with G. If $p \in X$, the functions arising from $\pi(p)$ are included in the set of functions arising from p. Conversely an inclusion of algebras of functions induces a map π which exhibits the maximal ideal space of the smaller algebra as a subflow of the maximal ideal space of the larger algebra.

A flow is *minimal* if every orbit is dense, i. e., if it contains no proper non-empty closed invariant subset. A function f in LUC is *right minimal* if for any T_α there is a T_β such that $T_\beta T_\alpha f = f$. If $f \in$ LUC, we denote by A_f the smallest Banach algebra containing f and its left translates, their conjugates, and the constants. If B is a self-adjoint Banach algebra in LUC containing the constants and invariant under left translation, we let $M(B)$ denote its maximal ideal space.

THEOREM 3-1. *A function $f \in$ LUC is right minimal if and only if $M(A_f)$ is a minimal flow. If B is a self-adjoint Banach algebra in* LUC *containing the constants and invariant under left translation, then the following conditions are equivalent:*

(1) *Every function in B is right minimal.*

(2) *B contains no proper non-zero self-adjoint ideals invariant under left translation.*

(3) *$M(B)$ is a minimal flow.*

(4) *$M(A_f)$ is a minimal flow for every $f \in B$.*

Proof. If f is right minimal, then so is every other function in A_f since the T_β corresponding to T_α can be chosen to be the same for all functions in A_f. Therefore the equivalence of (1) and (4) will imply the first statement of the theorem.

$(1) \Rightarrow (2)$. If (2) fails, let f be a function not identically zero in such an ideal, and let ρ be a non-zero multiplicative linear functional on LUC which vanishes on the ideal in B. Define a shift operator T_α by $T_\alpha h(g) = \rho(h_g)$. Then $T_\alpha f = 0$ and no T_β can recover f.

$(2) \Rightarrow (3)$. The closed invariant sets in $M(B)$ are in one-one correspondence with the closed self-adjoint left-invariant ideals in B.

$(3) \Rightarrow (4)$. A subflow of a minimal flow is minimal.

$(4) \Rightarrow (1)$. Let $M(A_f)$ be minimal and let T_α be given. Define a point $x \in M(A_f)$ by $x(h) = T_\alpha h(e)$, and by minimality of $M(A_f)$ find $\{g_n\}$ so that $g_n x$ converges to the evaluation-at-e functional. By compactness of the set of shift operators, choose a subnet of $\{g_n\}$ which defines some shift operator T_β. Then $T_\beta T_\alpha f = f$.

Functions arising from minimal flows were considered by Auslander and Hahn [1], who proved that (3) and (4) above are equivalent. They showed that for the additive group of reals the functions arising from minimal flows are not closed under addition.

LEMMA 3-2. *The uniform limit of right minimal functions is right minimal.*

Proof. Let $\lim f_n = f$ uniformly with each f_n right minimal, and let T_α be given. Choose T_{β_n} so that $T_{\beta_n} T_\alpha f_n = f_n$, and find a subnet of $\{T_{\beta_n}\}$ which converges, say to T_β. A 3ϵ argument then shows that $T_\beta T_\alpha f = f$.

THEOREM 3-3. *Any algebra A produced by Theorem* 1-1 *consists entirely of right minimal functions and is a maximal left-invariant self-adjoint algebra of right minimal functions. Conversely any left-invariant self-adjoint algebra of right minimal functions is contained in some A produced by that theorem. Consequently every maximal left-invariant self-adjoint algebra of right minimal functions is such an A, and the set-theoretic union of the A's is exactly the set of all right minimal functions.*

Proof. Since A and LUC$/I$ are isomorphic, A contains no proper non-zero self-adjoint left-invariant ideals and hence, by Theorem 3-1, consists entirely of right minimal functions. Any left-invariant self-adjoint algebra of right minimal functions clearly still consists of right minimal functions when the constants are adjoined. By Lemma 3-2 the uniform closure B of the resulting algebra contains only right minimal functions, and, by (2) of Theorem 3-1, condition (c) of Theorem 1-1 is satisfied by B. Thus B is contained in an algebra A produced by Theorem 1-1.

This result implies that every A produced by the theorem is maximal and that every maximal algebra is an A produced by the theorem. Finally we know that the union of the A's is contained in the set of right minimal functions. The reverse inclusion follows from the fact that if f is right minimal, then A_f is a left-invariant self-adjoint algebra of right minimal functions (Theorem 3-1) and is thus contained in some A.

The algebras A can be characterized quite simply in terms of shift operators. Call T_α minimal if $T_\alpha f$ is right minimal for every $f \in$ LUC; call T_α idempotent if $T_\alpha T_\alpha = T_\alpha$.

THEOREM 3-4. *The decompositions* LUC $= I \oplus A$ *of Theorem* 1-1 *stand* *in one-one correspondence with the minimal idempotent shift operators* T_u, *the correspondence being* $I=\ker T_u$ *and* $A=\text{image}\, T_u$.

Proof. If $\text{LUC}=I\oplus A$, put $\rho(i+a)=a(e)$ for $i\in I$ and $a\in A$, and define $T_uf(g)=\rho(f_g)$. Then T_u is an idempotent shift operator, and it is minimal by Theorem 3-3. Conversely if T_u is given, then LUC decomposes as $I\oplus A$ and the only problem is to show that I is maximal. But A contains no proper non-zero self-adjoint left-invariant ideals (Theorem 3-1) and hence the same is true of the isomorphic algebra LUC/I. Therefore I is maximal.

COROLLARY 3-5. *A function* f *in* LUC *is right minimal if and only if* $T_uf=f$ *for some minimal idempotent shift operator* T_u.

COROLLARY 3-6. *Any two maximal left-invariant self-adjoint algebras of right minimal functions are isomorphic.*

Proof. Let A_1 and A_2 be given, and put $\text{LUC}=I\oplus A_1=I\oplus A_2$ by Theorem 3-3 and Corollary 1-2. If T_u and T_v are the minimal idempotents corresponding to these decompositions, then the restriction of T_v to A_1 is the required isomorphism because the restriction of T_u to A_2 is a two-sided inverse.

A comparison of several classes of functions may be helpful. A right almost periodic function is a function in LUC whose right translates form a conditionally compact set in the uniform topology. Right almost periodic functions are also left almost periodic. If $f\in\text{LUC}$, f is right almost automorphic if any net $\{g_n\}$ in G contains a subnet $\{g_{n_j}\}$ such that $\lim_j f(gg_{n_j})$ exists and

$$\lim_k \lim_j f(gg_{n_k}{}^{-1}g_{n_j})=f(g).$$

These functions were considered by Bochner [2] and Veech [6]; Veech proved that every right almost automorphic function which is right uniformly continuous is left almost automorphic. If $f\in\text{LUC}$, f is right distal if an equality $T_\alpha T_\beta f=T_\alpha T_\gamma f$ always implies $T_\beta f=T_\gamma f$.

It is easy to see that right almost periodic functions are both right almost automorphic and right distal and that right almost automorphic functions are right minimal. Minimal idempotent shift operators T_u exist by Theorems 1-1 and 3-4, and hence if f is right distal the equality $T_uT_uf=T_uf$ gives $T_uf=f$. Hence right distal functions are right minimal. Using the results of Veech [6] and Bochner [2], one can show also that an almost automorphic right distal function is almost periodic.

We have seen that if f is right distal, then $T_uf=f$ for every minimal idempotent T_u. Veech's results imply that the same is true of all almost automorphic functions. Hence all almost automorphic functions, all right distal functions, and, in particular, all almost periodic functions are contained in every maximal self-adjoint left-invariant algebra of right minimal functions.

A right minimal function on a discrete group need not be left minimal. For an example, consider the action of the discretized unimodular matrix group $SL(2,R)$ on the circle, which is to be thought of as one-dimensional real projective space. This flow is transitive, hence minimal. If p is any point of the circle, then any function arising from the orbit of p is right minimal. On the other hand, each such function has a constant function as the limit of left translates and thus any non-constant such function cannot be left minimal.

We come now to the theorem which tells in what sense right minimal functions oscillate regularly. This result was suggested by H. Furstenberg. An element τ of G is a *right ϵ-translation number* for the function $f \in \mathrm{LUC}$ and for the subset F of G if $|f(g\tau)-f(g)|<\epsilon$ for all $g \in F$. A subset S of G is *right relatively dense* if there exist finitely many group elements $r_1, \cdots, r_n$ such that $G=\bigcup_{i=1}^{n} r_i S$.

For comparison, note that $f \in \mathrm{LUC}$ is right almost periodic if and only if the right ϵ-translation numbers for the set $F=G$ are right relatively dense.

THEOREM 3-7. *If $f \in \mathrm{LUC}$, then f is right minimal if and only if, for each finite subset F of G, the right ϵ-translation numbers for F are right relatively dense.*

Proof. Let f be right minimal and suppose that F is a finite set for which the set $\{\tau\}_F$ of right ϵ-translation numbers is not right relatively dense. To each finite set $E \subseteq G$, associate an element g_E not in $E\{\tau\}_F$; the result is a net in G indexed on the finite subsets of G. Choose a subnet $\{g_m\}$ which defines a shift operator T_α, and find T_β (defined by a net $\{h_n\}$) such that $T_\beta T_\alpha f=f$. Next fix h_n far enough out in the net that

$$|\lim_m f(gh_ng_m)-f(g)|<\epsilon/2$$

for all g in F, and choose an index m beyond the one-point set $\{h_n^{-1}\}$ for which

$$(*) \qquad |f(gh_ng_m)-f(g)|<\epsilon$$

for all g in F. On one hand, g_m was defined not to be in $h_n^{-1}\{\tau\}_F$, and on the other hand (*) gives $h_ng_m \in \{\tau\}_F$. Contradiction.

Conversely, let the right ϵ-translation numbers for each finite set F be right relatively dense and let T_α (defined by $\{g_m\}$) be given. We define a net $\{h_n\}$ on the directed set of pairs (F,ϵ). If (F,ϵ) is given, let S be the set of right ϵ-translation numbers for F and find finitely many r's so that $G = \cup\, rS$. Define $h_{(F,\epsilon)}$ to be an r^{-1} for which $\{g_m\}$ is in rS frequently. We claim that any shift operator T_β defined by a subnet of $\{h_n\}$ satisfies $T_\beta T_\alpha f = f$. It suffices to prove that if $g_0 \in G$ and $\delta > 0$ are given, then

$$| T_\alpha f(g_0 h_{(F,\epsilon)}) - f(g_0) | < \delta$$

whenever (F,ϵ) is beyond $(\{g_0\}, \delta/2)$. Thus choose such a pair (F,ϵ) and let m_0 be large enough that

$$| T_\alpha f(g_0 h_{(F,\epsilon)}) - f(g_0 h_{(F,\epsilon)} g_m) | < \delta/2$$

whenever $m \geqq m_0$. If S is the set of right ϵ-translation numbers for F, then g_m, by definition, is in $h_{(F,\epsilon)}{}^{-1}S$ frequently. Fix such $m \geqq m_0$. Then $h_{(F,\epsilon)} g_m$ is in S and

$$| f(g h_{(F,\epsilon)} g_m) - f(g) | < \epsilon \leqq \delta/2$$

for all $g \in F$ and in particular for $g = g_0$. Consequently

$$| T_\alpha f(g_0 h_{(F,\epsilon)}) - f(g_0) | < \delta.$$

4. Connection with the Ellis semigroup. We conclude by showing the connection between shift operators and the Ellis semigroup introduced in [4] and defined as follows. If (X, G) is a flow, we form the closure of G in the space X^X with the product topology. The result is a compact Hausdorff space which is a semigroup under composition and in which the maps $t \to ts$ for fixed s are continuous.

Let B be a self-adjoint Banach algebra in LUC containing the constants and invariant under left translation, and let $M(B)$ be the maximal ideal space of B. We shall use the symbol e interchangeably to denote the identity element of G and the evaluation-at-identity element of $M(B)$.

Relative to B we define an equivalence relation $\sim$ on the set of shift operators. Call $T_\alpha \sim T_\beta$ if $T_\alpha T_\gamma f = T_\beta T_\gamma f$ for all $f \in B$ and for all T_γ. Since the composition of two shift operators is a single shift operator, it follows that $T_\alpha \sim T_\beta$ implies $T_\epsilon T_\alpha \sim T_\epsilon T_\beta$ and $T_\alpha T_\epsilon \sim T_\beta T_\epsilon$. In addition, the relation $\sim$ is a closed relation since multiplication of shift operators is continuous in the first variable. In other words, if $\lim T_{\alpha_n} = T_\alpha$, $\lim T_{\beta_n} = T_\beta$, and $T_{\alpha_n} \sim T_{\beta_n}$ for all n, then $T_\alpha \sim T_\beta$. Therefore the quotient space of the set of shift operators modulo $\sim$ is a compact Hausdorff space with an inherited semigroup structure in which multiplication is continuous in the first variable. Convergence of a net of equivalence classes $[T_{\alpha_n}]$ to $[T_\alpha]$ means that $\lim T_{\alpha_n} T_\gamma f = T_\alpha T_\gamma f$ for all f in A and for all T_γ.

LEMMA 4-1. *Let B be a self-adjoint Banach algebra in* LUC *containing the constants and invariant under left translation, and let T_γ be a shift operator. Then there is a unique point x in $M(B)$ such that*

$$T_\gamma f(g) = f_g(x)$$

for all $g \in G$ and all $f \in B$. Every point x arises in this way from some T_γ.

Proof. The map $f \rightarrow T_\gamma f(e)$ is a multiplicative linear functional and hence establishes the existence and uniqueness of x. If x is given, choose a net of evaluations convergent to x and choose a subnet which defines a shift operator T_γ; then T_γ maps onto x.

THEOREM 4-2. *Let B be a self-adjoint Banach algebra in* LUC *containing the constants and invariant under left translation. Then the quotient of the set of shift operators by the equivalence relation induced by B is canonically topologically isomorphic to the Ellis semigroup $\bar{G}$ of $M(B)$.*

Proof. Let t be in $\bar{G}$, and put $Tf(g) = f_g(te)$ for $g \in G$ and $f \in B$. Let $\{g_n\}$ be any net of G-actions in $\bar{G}$ converging to t. Then $\lim f_g(g_n e) = f_g(te)$ by the continuity of f_g on $M(B)$, and hence

$$Tf(g) = \lim f(gg_n)$$

for all $f \in B$. Find a subnet of $\{g_n\}$ which defines a shift operator. We wish to show that different choices for these nets can lead only to equivalent shift operators. Thus let $\lim g_n = t$ in $\bar{G}$ with $\{g_n\}$ defining T_α and let $\lim h_m = t$ in $\bar{G}$ with $\{h_m\}$ defining T_β. Let T_γ be given and choose x as in Lemma 4-1. Then

$$(**) \qquad T_\alpha T_\gamma f(g) = \lim T_\gamma f(gg_n) = \lim f_{gg_n}(x) = \lim f_g(g_n x) = f_g(tx).$$

By symmetry $T_\alpha T_\gamma f = T_\beta T_\gamma f$, and we therefore have a well-defined map from $\bar{G}$ into the quotient space.

This map is one-one. In fact, if t and s are unequal elements of $\bar{G}$ corresponding to T_α and T_β, respectively, let $t(x) \neq s(x)$. If T_γ is chosen according to Lemma 4-1 so that it maps onto x and if f is a function which separates $t(x)$ and $s(x)$, then (**) shows that $T_\alpha T_\gamma(e) \neq T_\beta T_\gamma f(e)$. Hence T_α and T_γ are inequivalent.

The map is onto. For each x in $M(B)$ first choose T_γ by Lemma 4-1 such that $T_\gamma f(g) = f_g(x)$ and then find $t(x)$ by the same lemma such that

$$T_\alpha T_\gamma f(g) = f_g(t(x)).$$

Then t is in $M(A)^{M(A)}$. If $\{g_n\}$ is any net defining T_α, we claim that $\{g_n\}$ converges to t in $M(A)^{M(A)}$ and hence t is in $\bar{G}$. In fact, we have

$$f(t(x)) = T_\alpha T_\gamma f(e) = \lim T_\gamma f(g_n) = \lim f(g_n x).$$

If $\{g_n\}$ does not converge to t, then, by the compactness of $M(A)^{M(A)}$, some subnet converges to $s \neq t$ and gives the contradiction $f(t(x)) = f(s(x))$ for all f and x. Hence t is in $\bar{G}$, and (**) shows that t maps onto the equivalence class of T_α.

Finally we prove continuity. Let $\lim t_n = t$ in $\bar{G}$. We have $\lim f_g(t_n x) = f_g(tx)$ for all g, x, and f. Hence if t_n corresponds to $[T_{\alpha_n}]$ and t corresponds to $[T_\alpha]$, (**) shows that

$$\lim T_{\alpha_n} T_\gamma f(g) = T_\alpha T_\gamma f(g)$$

for all g, T_γ, and f. That is, $\lim [T_{\alpha_n}] = [T_\alpha]$. Thus the map is continuous and must be a homeomorphism.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

REFERENCES.

[1] L. Auslander and F. Hahn, "Real functions coming from flows on compact spaces and concepts of almost periodicity," *Transactions of the American Mathematical Society*, vol. 106 (1963), pp. 415-426.

[2] S. Bochner, "A new approach to almost periodicity," *Proceedings of the National Academy of Sciences*, vol. 48 (1962), pp. 2039-2043.

[3] R. Ellis, "Distal transformation groups," *Pacific Journal of Mathematics*, vol. 8 (1958), pp. 401-405.

[4] ———, "A semigroup associated with a transformation group," *Transactions of the American Mathematical Society*, vol. 94 (1960), pp. 272-281.

[5] A. W. Knapp, "Distal functions," *Proceedings of the National Academy of Sciences*, vol. 52 (1964), pp. 1409-1412.

[6] W. A. Veech, "Almost automorphic functions on groups," *American Journal of Mathematics*, vol. 87 (1965), pp. 719-751.

ETALE COVERINGS OF SCHEMES OVER HENSEL RINGS.

By M. Artin.*

Let A be a henselian local ring [9] and X a prescheme over Spec A. The object of this paper is to study the relation between etale coverings of X and etale coverings of the prescheme $\hat{X}$ deduced from X by the base change $A \to \hat{A}$ (where $\hat{A}$ is the completion of A). Our main results (II, (1, 3), (2. 1), (4. 2)) assert that in certain cases the natural functor $Et(X) \to Et(\hat{X})$ relating the categories of these coverings is an equivalence, so that a covering of $\hat{X}$ is obtained in exactly one way, up to canonical isomorphism, from a covering of X. If X is proper over Spec A, this follows immediately from the theorem of specialization of the fundamental group ([1], XII (5. 9)), which asserts that the restriction from X to its closed fibre X_0 induces an equivalence on etale coverings. Thus we are primarily interested in the case that X is not proper, for instance that X is an open subset of Spec A. We discuss this last case in Section 5 of part II. However, our proofs make heavy use of the theorem for a proper X.

In [10], we announced some of our results and showed how they could be used to prove that a complete normal 2-dimensional local ring of residue characteristic zero is *algebraic*, i. e., is the completion of a local ring of an algebraic surface. The method could probably be pushed through to give the same result for isolated singularities of dimension 3. In the meantime, Hironaka (to appear) has developed an elegant technique which gives a somewhat more general result in this direction and which works in characteristic p. As far as I know, the conjecture of Grauert that any isolated singularity is algebraic is still open in the general case.

We use freely a number of techniques of Grothendieck, such as descent theory ([5], VIII, IX) and local Lefschetz theory ([6], X), and we assume the reader familiar with the notions of henselization and strict localization (cf. [9] and [1], VIII), and of excellent ring ([4], IV, 7. 8). Some auxiliary material is collected in part I. The main discussion starts in part II.

Received November 15, 1965.

* This research was partially supported by NSF GP 1740.

Part I. Auxiliary results.

1. Approximation of equations in Hensel Rings.

Definition (1.1). Let A be a ring and $\mathfrak{A}$ an ideal of A. The pair $(A,\mathfrak{A})$ is called *henselian* iff. given $f(x), g(x) \in A[x]$ and $a \in A$ satisfying

$$f(x) \equiv (x-a)g(x) \pmod{\mathfrak{A}}$$

and

$$f'(a) = g(a) \equiv 1 \pmod{\mathfrak{A}}$$

there exist $g_1(x) \in A[x]$ and $a_1 \in A$ such that

$$f(x) = (x-a_1)g_1(x)$$

and

$$a \equiv a_1 \pmod{\mathfrak{A}}.$$

A local ring is called henselian iff. the pair $(A,\mathfrak{M})$ is henselian, where $\mathfrak{M}$ is the maximal ideal (cf. [9]).

The lemma below is a generalization of ([2] § 4, No. 5, Cor. 1) and the method of proof is similar. I am indebted to J-P. Serre for pointing out that my original argument could be greatly simplified.

LEMMA (1.2). *Let $(A,\mathfrak{A})$ be henselian and $f(x) = (x-a)g(x)$ an equation in $A[x]$. Put $b = g(a)$. If $f_1(x) \in A[x]$ is a polynomial such that*

$$f_1(x) \equiv f(x) \pmod{b^2\mathfrak{A}}$$

then there exists a_1 and $g_1(x)$ such that

$$f_1(x) = (x-a_1)g_1(x), \text{ and}$$

$$a \equiv a_1 \pmod{b\mathfrak{A}}.$$

Proof. It is clearly permissible to make a linear substitution on x so that $a = 0$. Then the lemma reads: If

$$(*) \qquad f_1(x) \equiv xg(x) \pmod{b^2\mathfrak{A}},\ b = g(0),$$

then there exists a root ϵ of $f_1(x)$ such that

$$\epsilon \equiv 0 \pmod{b\mathfrak{A}}.$$

To prove this, consider the substitution $bu = x$. It suffices to find a root δ of the polynomial $f_1(bu) \in A[u]$, with

$$\delta \equiv 0 \pmod{\mathfrak{A}}.$$

Since $b=g(0)$, the polynomial $g(bu)$ is divisible by b. Therefore

$$bug(bu)=b^2u\tilde{g}(u)$$

for some $\tilde{g}\in A[u]$ satisfying $\tilde{g}(0)=1$. Therefore by (*),

$$f_1(bu)=b^2\bar{f}(u)$$

with

$$\bar{f}(u)\equiv u\tilde{g}(u) \qquad (\operatorname{mod}\mathfrak{A}),$$

and

$$\bar{f}'(0)=\tilde{g}(0)=1.$$

Since $(A,\mathfrak{A})$ is henselian, $\bar{f}$, and hence f_1, has a root $\delta\equiv 0$ $(\operatorname{mod}\mathfrak{A})$.

PROPOSITION (1.3). *Let A be a henselian local ring, and $f(x)\in A[x]$ a monic polynomial. Suppose a monic $f_1(x)\in A[x]$ satisfies*

$$f_1(x)\equiv f(x) \qquad (\operatorname{mod}\delta^2\mathfrak{m})$$

where δ is the discriminant of f and $\mathfrak{m}=\max A$. *Then there is an A-isomorphism*

$$A[x]/(f)\xrightarrow{\sim}A[x]/(f_1)$$

congruent to the identity $(\operatorname{mod}\mathfrak{m})$.

Proof. Put $B=A[x]/(f)$, $B_1=A[x]/(f_1)$. If we find a root α_1 of $f_1(x)$ in B with α_1 congruent $(\operatorname{mod}\mathfrak{m}B)$ to the residue $\bar{x}$ of x, we can map $B_1\to B$ by $\bar{x}_1\rightsquigarrow\alpha_1$. This will be an isomorphism by the Nakayama lemma. Now B is a finite A-algebra, hence a direct product of henselian local rings [9], and so it is immediate that $(B,\mathfrak{m}B)$ is henselian. Thus we may apply Proposition (1.2), with $(B,\mathfrak{m}B)=(A,\mathfrak{A})$.

PROPOSITION (1.4). *Let A be an excellent hensel ring, and let B, C be finite A-algebras. Assume C reduced. Denote by $\hat{X}$ the completion of X. Then the natural map*

$$\operatorname{Hom}_A(B,C)\to\operatorname{Hom}_{\hat{A}}(\hat{B},\hat{C})$$

is bijective.

Proof. Since $C\to\hat{C}$ is injective, the map is injective. We may assume C normal (in the general sense, i.e., a product of normal integral domains). For, let $\bar{C}$ be the normalization of C, and suppose the result known for $\bar{C}$. Then $C=\hat{C}\cap\bar{C}$ (this is true for any finite A-modules $M\subset N$). If a map $\hat{B}\to\hat{C}$ is given, it induces a map $B\to\hat{C}$, hence $B\to\bar{C}$. The image of B is in $C \cap \bar{C} = C$, hence we get a map $B \rightarrow C$ as desired. Thus we are reduced to the case C normal. We may also suppose C entire, i. e., an integral domain.

Now a map $B \rightarrow C$ is equivalent with a section of $B \otimes_A C$ over C. Replacing A by C and B by $B \otimes_A C$, we are reduced to showing that if A is normal and entire then sections of B over A are in one to one correspondence with section of $\hat{B}$ over $\hat{A}$. We may assume B reduced. Since the minimal primes of B and $\hat{B}$ correspond ([4], IV. 7. 8 (vii)), we may assume B integral. Then the assertion is merely $B = A$ iff. $B = \hat{A}$. This is true by the Nakayama lemma.

2. Some remarks on geometrically unibranch schemes. Recall that according to Grothendieck ([4], IV. 6. 15) a prescheme X is *geometrically unibranch* iff. every local ring A of X has the following property:

(2.1) The reduced ring A_{red} is entire, the integral closure $\bar{A}_{\mathrm{red}}$ of A_{red} is a local ring, and the residue field of $\bar{A}_{\mathrm{red}}$ is purely inseparable over the residue field of A.

An equivalent definition is the following:

LEMMA (2.2). *A prescheme X is geometrically unibranch iff. for every geometric point $\bar{x}$ of X the strict localization $\tilde{A}$ of X at $\bar{x}$ has the property that $\tilde{A}_{\mathrm{red}}$ is entire.*

Proof. Let x be the image of $\bar{x}$ in X. The strictly local ring $\tilde{A}$ is the *strict localization* of the local ring A of X at x, i. e., the maximal unramified extension of the henselization of A ([1], VIII). This strict localization is a reduced ring iff. A is, and reduction commutes with strict localization. Therefore what we need to show is that (2.1) for a reduced ring A is equivalent with he assertion that the strict localization $\tilde{A}$ of A is entire. We may suppose that A is entire. Suppose first that (2.1) does not hold for A. Then since the normalization $\bar{A}$ is a limit of finite subextensions of A, there exists a finite A-algebra $A_1 \subset \bar{A}$ such that either (i) A_1 is not local, or (ii) the residue field of A_1 contains a separable extension of that of A. Let $\tilde{A}_1 = A_1 \otimes_A \tilde{A}$. Since $\tilde{A}$ is henselian and $\tilde{A}_1$ is a finite $\tilde{A}$-algeba, $\tilde{A}_1$ is a product of local rings. But it is clear that by construction $\tilde{A}_1$ is not local (in case (ii) because $\tilde{A}$ has separably closed residue field). Therefore $\tilde{A}_1$ is not entire. However, if $\tilde{A}$ were entire, $\tilde{A}_1$ would be an integal extension. Thus $\tilde{A}$ is not entire either.

Conversely, if $\tilde{A}$ is entire, then the above argument in reverse shows that any finite A-algebra $A_1 \subset \bar{A}$ must be local and have residue field radicial over A. Since $\tilde{A}$ is an inductive limit of such algebras, the same holds for $\tilde{A}$, and thus (2.1) is satisfied.

An immediate consequence of Hartshorne's connectedness theorem ([7]; [6], 3, III-1) is the following criterion (analogous to Serre's normality criterion, but of course not a necessary condition):

PROPOSITION (2.3) (Hartshorne's unibranch condition). *A locally noetherian prescheme X is geometrically unibranch if it has the following properties:*

(i) *It is geometrically unibranch in codimension* 1, *i.e., for every point $x \in X$ of codimension* 1, *the local ring A of X at x satisfies* (2.1).

(ii) *For every point $x \in X$ of codimension $\geqq 2$, the local ring A of X at x has depth $\geqq 2$.*

Proof. Let A be a local ring of X. We need to show that (2.1) holds for A. By induction on $\dim A$, we may suppose every point of $\operatorname{Spec} A$ outside the closed point does satify (2.1). If $\dim A = 1$, we are through by property (i). Suppose $\dim A \geqq 2$ (so that depth $A \geqq 2$ by (ii)). Let $\tilde{A}$ be the strict localization of A. $\tilde{A}$ is a noetherian ring. For any prime $\tilde{p}$ of $\tilde{A}$, let $p = \tilde{p} \cap A$. Since $\tilde{A}$ is a limit of etale extensions of A, it is immediately seen that A_p and $\tilde{A}_{\tilde{p}}$ have isomorphic strict localizations. Therefore every point of $\operatorname{Spec} \tilde{A} - \{a\}$ (a the closed point) satisfies (2.1). Since $\tilde{A}$ is faithfully flat over A, depth $\tilde{A} \geqq 2$. Therefore $\operatorname{Spec} \tilde{A} - \{a\}$ is connected by Hartsthorne's connectedness theorem [7]. If $\tilde{A}_{\mathrm{red}}$ were not entire, there would be a prime $\tilde{p}$ of $\tilde{A}$, $\tilde{p} \neq a$, in common to two irreducible components of $\operatorname{Spec} \tilde{A}_{\mathrm{red}}$, and this would contradict (2.1) for $\tilde{p}$. Thus $\tilde{A}_{\mathrm{red}}$ is entire, and hence (2.1) holds for A.

LEMMA (2.4). *Let A be an excellent local ring. Then* $\operatorname{Spec} A$ *is geometrically unibranch iff.* $\operatorname{Spec} \hat{A}$ *is.*

Proof. We may replace A by A_{red} and hence suppose A reduced, in fact entire. The normalization $\bar{A}$ of A is a finite A-module ([4], IV. 7. 8. 3 (vi)), hence $\bar{A} = \hat{A} \otimes_A \bar{A}$ is again normal ([4], IV. 7. 8. 3 (v)). By [4], 0_{IV}. 23. 2. 2., A is geometrically unibranch iff. the map $\operatorname{Spec} \bar{A} \rightarrow \operatorname{Spec} A$ is radicial, which is true iff. $\operatorname{Spec} \hat{\bar{A}} \rightarrow \operatorname{Spec} \hat{A}$ is radicial ([5], VIII. 3. 1), i.e., iff. $\operatorname{Spec} \hat{A}$ is geometrically unibranch.

Using Hartshorne's condition (2.3), we can prove the following result on "ground field change" in complete local rings:

THEOREM (2.5). *Let A be a complete local ring and $k \subset A$ a field such* *that* $A/\mathfrak{m}A$ *is finite over* k. *Let* $L \supset k$ *be a field extension and denote by* A_L *the completion of* $A \otimes_k L$ *with respect to the ideal generated by* $\mathfrak{m}$. *Let* $X/\operatorname{Spec} A$ *be a prescheme of finite type which is geometrically unibranch. Then* $X_L = X \times_{\operatorname{Spec} A} \operatorname{Spec} A_L$ *is again geometrically unibranch.*

Proof. Let $\bar{X}$ (resp. $\bar{X}_L$) be the normalization, in the general sense, of X (resp. X_L). (By normalization of a noetherian ring A we mean the product of the normalizations of the rings A/p_i, where $\{p_i\}$ are the minimal prime ideals.) We have canonical maps $\bar{X}_L \xrightarrow{a} X_L \times_X \bar{X} \xrightarrow{b} X_L$. By assumption, the map $\bar{X} \to X$ is radicial, hence b is, and we need to show that ba is radicial ([4], 0_{IV}. 23. 2. 2). Therefore it suffices to show a radicial. Since A is an excellent ring, $\bar{X}$ is again of finite type over $\operatorname{Spec} A$. Thus we are reduced to the case $X = \bar{X}$, i. e., X normal. The problem is local on X.

We will verify the conditions of (2. 3). Let pX_L be a point of codimension $\geqq 2$, and let p_0 be its image in X. If p_0 has codimension $\geqq 2$, then $\operatorname{depth}_{p_0} X \geqq 2$ since X is normal, hence $\operatorname{depth}_p X_L \geqq 2$ since X_L is flat over X. Say $\operatorname{codim}_{p_0} X = 1$. Choose a function t on Y which is a local parameter at the point p_0, and let $Y = V(t)$ be the locus $\{t = 0\}$. Since t is a nonzero-divisor at p_0, it suffices to show that $\operatorname{depth}_p Y_L \geqq 1$, hence it suffices to prove the following assertion:

If $p \in Y_L$ lies over the generic point of an irreducible scheme Y, then $\operatorname{depth}_p Y_L = \operatorname{codim}_p Y_L$.

This will also handle the case that $\operatorname{codim}_{p_0} X = 0$. To prove it, note that we may assume, replacing A by a quotient if necessary, that the map $Y \to \operatorname{Spec} A$ is generically surjective. Since $A/\mathfrak{m}A$ is finite over k, A is a finite algebra over a power series ring $k[[x_1, \cdots, x_n]]$ ($n = \dim A$). Hence we may assume $A = k[[x_1, \cdots, x_n]]$. Next, if d is the dimension of the fibre of $Y/\operatorname{Spec} A$ at p_0, Y will have a neighborhood of p_0 which admits a finite generically surjective map to an affine space $\operatorname{Spec} A[y_1, \cdots, y_d]$ over A. Since p_0 is the generic point of Y, this map is flat at p_0. Therefore, depth and codimension being preserved in flat finite extensions, we are reduced to the case $Y = \operatorname{Spec} A[y_1, \cdots, y_d]$. Then $Y_L = \operatorname{Spec} A_L[y_1, \cdots, y_d]$ where

$$A_L = L[[x_1, \cdots, x_n]].$$

Thus Y_L is regular, and so the assertion follows.

It remains to check condition (i) of (2. 3) for the scheme X_L. Let $p \in X_L$ have codimensional 1. If p lies over a point p_0 of X of codimension zero, then the construction used above for Y shows what we want. In fact, locally at p_0 there is a finite map $X \to \operatorname{Spec} A[y_1, \cdots, y_d]$ which can be factored into a radicial map $X \to X'$ followed by an etale map $X' \to \operatorname{Spec} A[y]$. Now it is immediate that if $A' \to A$ is a finite radicial extension of local rings of dimension 1 and if A' is geometrically unibranch, then A is too. Since X'_L is regular, it follows that X_L is geometrically unibranch at p.

Suppose now that the image p_0 of p in X has codimension 1. Since the problem is local on X, we may suppose X a normal projective variety. Then there is a point q which is a specialization of p such that its image q_0 in X is a closed point of the closed fibre of $X/\operatorname{Spec} A$. Let B be the local ring of X at q_0, and B' the local ring of X_L at q. It suffices to show that $\operatorname{Spec} B'$ is geometrically unibranch, hence to show that $\operatorname{Spec} \hat{B}'$ is geometrically unibranch by (2.4). But clearly $\hat{B}' = \hat{B}_L$ is the completion of $B \otimes_k L$. Thus we are reduced to the case $X = \operatorname{Spec} A$ with A a normal ring, and to showing that X_L is geometrically unibranch at a point p of codimension 1 whose image p_0 is also of codimension 1. Choose a system of parameters $x_1, \cdots, x_n$ of A such that x_n is zero at p_0. Then in the finite map $\operatorname{Spec} A \to \operatorname{Spec} k[\![x_1, \cdots, x_n]\!]$, p_0 is in the inverse image of the locus $\{x_n = 0\}$. Let C be the local ring of $k[\![x_1, \cdots, x_n]\!] = A_0$ at the general point of $\{x_n = 0\}$, $B = C \otimes_{A_0} A$, C' the local ring of $L[\![x_1, \cdots, x_n]\!]$ at the general point of $\{x_n = 0\}$, and $B' = C' \otimes_C B$. We must prove that B' is geometrically unibranch, hence that its completion B' is geometrically unibranch (2.4). Set

$$k' = k((x_1, \cdots, x_{n-1})), \qquad L' = L((x_1, \cdots, x_n)).$$

Then clearly $\hat{C}'$ is the completion $\hat{C}_{L'}$ of $\hat{C} \otimes_{k'} L'$, hence B' is the completion $\hat{B}_{L'}$ of $\hat{B} \otimes_k L'$. Thus we are reduced to the case that A is a complete semilocal ring of dimension 1, and to prove that if A is geometrically unibranch then A_L is. It is clear from (2.2) that it is permissible to replace k be a finite field extension k' and A by $A \otimes_k k'$. For, a separable extension does not effect (2.2), nor does a finite radicial one. Thus we may assume that A is local and that k is separably algebraically closed in $A/\mathfrak{m}A$. We may moreover assume A entire. Then (2.2) implies that A is geometrically unibranch iff. k is separably closed in the field of fractions K of A. We have to show that A_L is ireducible and that L is separably closed in the field of fractions K_L of $(A_L)_{\mathrm{red}}$. If L is not separably closed in K_L, then replacing L by a finite separable extension L' results in a situation where $A_{L'} = A_L \otimes_L L'$ is not irreducible. Thus it suffices to show that A_L is irreducible. Consider a subring $k[\![t]\!]$ of A such that A is finite over $k[\![t]\!]$. Then $A_L = A \otimes_{k[[t]]} L[\![t]\!]$. Hence $K_L = (K \otimes_{k((t))} L((t)))_{\mathrm{red}}$. If K_L is not a field, there is an element of K separable over $k((t))$ such that its irreducible equation $f(x) = 0$ has a root in $L((t))$. Since by assumption k is separably closed in K, it suffices to show the following:

Let $k \subset L$ be fields and α an element of $L((t))$ which is separably algebraic over $k((t))$. Then α lies in $k'((t))$ for some finite extension k' of k.

But this is a well known and easy consequence of the continuity of roots of equations in $k((t))$. Thus the proof is complete.

3. A lemma on extension of coefficient fields. The result is the following:

Proposition (3.1). *Let $A \rightarrow B$ be a local homomorphism of complete local rings and let $k \subset A$ be a coefficient field. Suppose $\bar{B} = B/\mathfrak{m}B$ ($\mathfrak{m} = \max A$) is regular over $k = A/\mathfrak{m}$, i. e., that $\bar{B} \otimes_k k'$ is a regular ring for every finite extension field k' of k. There is a finite radicial extension k' of k such that, setting $B' = B \otimes_k k'$, $A' = A \otimes_k k'$, $\mathfrak{m}' = \max A'$, $\bar{B}' = B'/\mathfrak{m}B'$, the following holds: There exist elements $z_1, \cdots, z_r \in B'$ which reduce to part of a regular system of parameters for $\bar{B}'$ such that k' extends to a coefficient field of $B'/(z_1, \cdots, z_r)$. In characteristic $p \neq 0$, the extension k' may be chosen so as to be obtained by adjoining p-th roots of some members of a p-base for k given in advance.*

Proof. The result is trivial in characteristic zero. Assume that the characteristic is $p \neq 0$. Let L be the residue field of B. For any finite set of elements $S = \{a_1, \cdots, a_N\}$ of k, put

$$\delta(S) = p\text{-rank}_k(S) - p\text{-rank}_L(S).$$

The number $\delta(S)$ is bounded by $\dim \bar{B}$. For, if $\{a_1, \cdots, a_N\}$ is a set, say p-independent over k, and if $\{a_1, \cdots, a_r\}$ is a maximal p-independent subset over L, let k' be the field obtained by adjoining the p-th roots of $\{a_1, \cdots, a_r\}$ to k. Then $\bar{B}' = \bar{B} \otimes_k k'$ is regular with residue field $L'[\{a_i^{1/p}\}]$. Now replacing k by k', we have the same situation except that the p-independent set $\{a_{r+1}, \cdots, a_N\}$ over k is now actually in L^p. We may therefore assume that this is so from the start. Let k' denote the field obtained by adjoining the p-th roots of $\{a_{r+1}, \cdots, a_N\}$ to k. Then B' is regular with residue field $L' = L$, and $\dim \bar{B}' = \dim \bar{B}$. But $L \otimes_k k'$ is a quotient of $\bar{B}'$, and since $a_{r+1}, \cdots, a_N$ are p-th powers in L,

$$L \otimes_k k' \approx L[\epsilon_{r+1}, \cdots, \epsilon_N]$$

where $\epsilon_i{}^p = 0$. Thus necessarily $n - r \leqq \dim \bar{B}'$, which proves the assertion.

Let therefore the maximum for $\delta(S)$ be d. Replacing k by a radicial extension as above, we may suppose that there are elements $a_1, \cdots, a_d \in k$ which are p-independent over k but which are in L^p. It is easily seen that $a_1, \cdots, a_d$ may be chosen to be part of a given p-basis for k, and that k'

may be obtained by extracting p-th roots of elements of that basis. Obviously the following is true: If a set $\{a_1, \cdots, a_d; b_\nu\}$ of elements of k is p-independent over k, then $\{b_\nu\}$ is p-independent over L.

Now choose a p-basis $\{a_1, \cdots, a_d; b_\nu\}$ for k, and extend the set $\{b_\nu\}$ to a p-basis $\{b_\nu, c_\mu\}$ for L. Choose representatives for the elements c_μ in B and denote them by the same letter (the elements b_ν are already canonically represented in B). This set of representatives determines a coefficient field, say K, of B.

Let $a_1', \cdots, a_d' \in K$ be the elements whose residues in L are $a_1, \cdots, a_d$. Actually, $a_i'^p = a_i'$ $(i = 1, \cdots, d)$. Consider the extension $k' = k[\alpha_1, \cdots, \alpha_d]$ of k, where $\alpha_i^p = a_i$. Put $a_i - a_i' = \xi_i \in \max B$. Since $a_i' \in K^p$, the extension $B' = B \otimes_k k'$ is isomorphic to

$$B[y_1, \cdots, y_d], \text{ where } y_i^p = \xi_i.$$

Now $\bar{B}' = \bar{B} \otimes_k k'$ is regular by assumption, hence the elements $Y_i^p - \xi_i$, $i = 1, \cdots, d$, form part of a regular system of parameters at the origin for the ring $\bar{B}[Y_1, \cdots, Y_d]$ (we denote also by ξ_i the images in $\bar{B}$). Therefore the elements ξ_i do too, and hence they are part of a system of parameters for $\bar{B}$. Now in $C = B/(\xi_1, \cdots, \xi_d)$, the residues of a_i and a_i' are equal. Hence the p-basis $\{a_i, b_\nu\}$ of k is mapped into the coefficient field K of C. It is easily seen that this implies that k is mapped into K. This completes the proof.

Part II. The main theorems.

1. The case of characteristic zero. In this part, we make free use of descent theory ([5], VIII, IX) and of ([1], XII. (5.9)). Given a prescheme Z, we denote by $Et(Z)$ the category of finite etale coverings of Z.

PROPOSITION (1.1).[1] *Let A be an excellent hensel ring, and X a prescheme over* Spec A. *Let ^ denote the base change from A to $\hat{A}$. The canonical functor*

$$Et(X) \to Et(\hat{X})$$

is fully faithful.

Proof. We need to show that if U, V are objects of $Et(X)$, then the map $\mathrm{Hom}_X(U, V) \to \mathrm{Hom}_{\hat{X}}(\hat{U}, \hat{V})$ is bijective. The map is clearly injective since $A \to \hat{A}$ is faithfully flat.

We may suppose X affine. For, let $\{X_i\}$ be an affine covsring of X, and put $X' = \coprod_i X_i$, $X'' = \coprod_{i,j} X_i \cap X_j$. Then denoting by primes the base changes, we have a diagram

---

[1] A more general version of this result is in [4] Ch. IV.

(1.2)
$$\begin{tikzcd}
\mathrm{Hom}_X(U, V) \arrow[r] \arrow[d, "a"] & \mathrm{Hom}_{X'}(U', V') \arrow[r, shift left] \arrow[r, shift right] \arrow[d, "b"] & \mathrm{Hom}_{X''}(U'', V'') \arrow[d, "c"] \\
\mathrm{Hom}_{\hat X}(\hat U, \hat V) \arrow[r] & \mathrm{Hom}_{\hat X'}(\hat U', \hat V') \arrow[r, shift left] \arrow[r, shift right] & \mathrm{Hom}_{\hat X''}(\hat U'', \hat V'')
\end{tikzcd}$$

where the horizontal rows are exact and the vertical arrows are injective. If the result is true for an affine, it follows that b is bijective because $\mathrm{Hom}_{X'}(U', V') = \prod_i \mathrm{Hom}_{X_i}(U_i, V_i)$. This together with the injectivity of c implies that a is bijective, as required.

Now an affine scheme $X/\mathrm{Spec}\,A$ is an inverse limit of affines which are of finite type over A. Since the limit commutes with base extension, a standard argument of passage to the limit, of the type of [4], IV. 8 reduces one to the case that X is of finite type. We omit this argument.

Any proper and surjective map is an effective morphism of descent for the category of etale coverings ([5], IX). Therefore it suffices to prove the theorem for any X' proper and surjective over X. Hence we may assume X normal and affine.

Let U, V be given. The graph of a map $\hat U \to \hat V$ over $\hat X$ is a connected component of $\hat U \times_{\hat X} \hat V$, since U, V are affine and hence separated over X. Replacing X by U and V by $U \times_X V$, we are reduced to showing that the sections of V/X and $\hat V/\hat X$ correspond, i. e., that V is connected iff. $\hat V$ is connected. Now V is affine and normal, and hence extends to a normal scheme $\bar V$ projective over $\mathrm{Spec}\,A$. Thus we are reduced to showing that for a proper scheme $\bar V/\mathrm{Spec}\,A$, $\bar V$ is connected iff. $\hat{\bar V}$ is connected. This is an easy consequence of hensel's lemma in A, and may be found in ([1], XIII (5.7)). In fact, the connected components of each scheme correspond to those of their closed fibres, and their closed fibres are equal.

THEOREM (1.3). *Let A be an excellent hensel ring with residue characteristic zero, and let $X/\mathrm{Spec}\,A$ be a prescheme. The functor $Et(X) \to Et(\hat X)$ is an equivalence of categories.*

Proof. By (1.1) the functor is fully faithful. Thus it remains to show that every finite etale covering $\hat U$ of $\hat X$ is induced up to isomorphism by a covering U of X. Let $X' \to X$ be an effective morphism of descent for the fibred category $Et([5], IX)$, and let $\hat U$ be a finite etale covering of X. If $\hat U' = \hat U \times_{\hat X} \hat X'$ is induced by a covering U' of X', then the fact that the functors $Et(X'') \to Et(\hat X'')$ (resp. $Et(X''') \to Et(\hat X'''))$ are fully faithful by (1.1) $(X'' = X' \times_X X')$ implies that canonical descent data for the covering U' and the map $X' \to X$ are given, hence that $\hat U$ is induced by an etale covering

U of X. Thus we may replace X by any X'/X of this type. Replacing X by an affine cover shows that it suffices to consider the case X affine. Then a limit argument of the usual type ([4], IV. 8) shows that X may be taken of finite type.

The proof we give here uses resolution of singularities [8]. Another (longer) proof results from the methods of No. 4 (cf. (5.1)), which would not be simplified by use of resolution: Since X is affine and of finite type over Spec A, it is an open subscheme of some $\bar{X}$ projective over Spec A. The proper base change theorem ([1], XII (5.9)) implies that the categories $Et(\bar{X})$ and $Et(\hat{\bar{X}})$ are equivalent, namely that they are both equivalent to $Et(\bar{X}_0)$ where $\bar{X}_0$ is the closed fibre of $X/\mathrm{Spec}\, A$ (resp. $\hat{X}/\mathrm{Spec}\, \hat{A}$). Thus if our given covering $\hat{U}/\hat{X}$ extends to a covering $\hat{\bar{U}}$ of $\hat{\bar{X}}$, we are through. Now it is still permissible to make a proper base change $\bar{X}' \to \bar{X}$ since proper morphisms are effective morphisms of descent for the fibred category Et. Therefore by resolution of singularities [8] we may assume that $\bar{X}$ is integral and nonsingular and that $Y = \bar{X} - X$ has only normal crossings. Then since A is excellent, $\hat{\bar{X}}$ is also nonsingular and $\hat{Y}$ has only the normal crossings induced by those of Y.

Let P be an arbitrary point of $\bar{X}$, let Y_i $(i = 1, \cdots, n)$ be the components of Y passing through P, and let r_i be a number dividing the ramification indices of the extension $\hat{U}$ of $\hat{X}$ along $\hat{Y}_i$. Let f_i be a rational function on X which is a local equation for Y_i at P, and let $\bar{X}'$ (resp. X') be obtained from $\bar{X}$ (resp. X) by adjoining the r_i-th roots of f_i $(i = 1, \cdots, n)$. Then $\bar{X}'$ is nonsingular above P, and by Abyankhar's lemma ([5], X. 3. 6), the extension $\hat{U}'$ of $\hat{X}'$ induced by $\hat{U}/\hat{X}$ is unramified along any divisor above P, hence unramified by the theorem of purity of the branch locus ([5], X. 3. 1). Therefore the normalization $\hat{\bar{U}}'$ of $\hat{\bar{X}}'$ in the ring of rational functions $R(\hat{U}')$ of $\hat{U}'$ is etale at P. Thus for any P a proper surjective map $\bar{X}' \to \bar{X}$ can be found such that the induced covering $\hat{U}'$ extends to an etale covering of $\hat{\bar{X}}'$ above P. By compactness, there exist a finite number of such maps, say $\bar{X}_i \to \bar{X}$ $(i = 1, \cdots, N)$ such that one of them works for any given P. Let $\bar{X}'$ be the normalization of the fibred product of the $\bar{X}_i$ over $\bar{X}$. It is immediate that the normalization $\hat{\bar{U}}'$ of $\hat{\bar{X}}'$ in $R(\hat{U}')$ $(U' = U \times_X X')$ is etale everywhere over $\hat{\bar{X}}'$. This completes the proof.

2. Removal of the closed fibre of a proper map.

THEOREM (2.1). *Let $A \subset A^*$ be excellent hensel rings such that the inclusion induces an isomorphism of their completions. Let $X/\mathrm{Spec}\, A$ be proper, $Y \subset X$ a closed subscheme, and $U = X - Y$. Let a star denote the base change from A to A^*. Suppose that the map $Y \leftarrow Y^*$ is an isomorphism* *(for instance, that Y lies above the closed point of* Spec A*). Then the functor* $Et(U) \to Et(U^*)$ *is an equivalence of categories.*

Proof. It follows immediately from (1.1) that the functor is fully faithful for any scheme over Spec A. Thus it remains to show that a given finite etale covering V^*/U^* is induced (up to isomorphism) by an etale covering V/U. If V^* extends to an etale covering $\bar{V}^*$ of X^*, then the proper base change theorem ([1], XII(5.9)) implies that $\bar{V}^*$ is induced by an etale covering of X, which yields the result (cf. proof of (1.3)). Moreover, if $X' \to X$ is a proper surjective map, and if the covering $V'^* = V^* \times_{X^*} X'^*$ of U'^* is induced by an etale covering V' of U', then the fact that the functor is fully faithful implies that we are given canonical descent data for V'/X' adn for the map $U' \to U$, hence that V^* is induced by a covering V of U (cf. proof of (1.3)). Letting $Y' = Y \times_X X'$, we have $Y' \xleftarrow{\sim} Y'^*$. Hence the hypotheses of the theorem are preserved under a proper base change, and so we may replace X by any X' projective and surjective over it. We may in particular take X to be irreducible and normal.

Let us try to find such an X'/X so that the covering V'^* extends to an etale covering $\bar{V}'^*$ of X'^* (this will complete the proof). It is a local problem on X. For, if locally on X such a projective map $X' \to X$ is found, then, say taking X' normal, X' extends to a scheme surjective and projective over all of X. If $\{X_i \to X\}$ is a finite set of these surjective projective maps such that the covering extends over a given $P \in X$ for at least one X_i, then the normalization X' of the product of the X_i over X clearly has the property that the normalization $\bar{V}'^*$ of X'^* in the ring of rational functions $R(V'^*)$ is etale, i.e., that V'^* extends to an etale covering of X'^*.

Actually, the problem is also local for the etale topology. In fact, let P be a point of X, $X_1 \to X$ an etale map, P_1 a point of X_1 lying over P, and $X_1' \to X_1$ a surjective projective map such that the covering $V_1'^*$ of $U_1'^*$ extends to an etale covering $\bar{V}_1'^*$ of X_1'. Let K (resp. K_1) be the function field of X (resp. X_1 (taking X_1 connected)) and let K_2/K be a galois extension containing K_1, say with group G. Let $\bar{X}_i$ be the normalization of X in K_i $(i = 1, 2)$. The scheme X_1' extends to a scheme $\bar{X}_i'$ projective over $\bar{X}_1$. Put $\bar{X} = X_2 \times_{X_1} \bar{X}_i'$, and for $\sigma \in G$ denote by $\bar{X}^\sigma$ the scheme obtained by translating $\bar{X}$ by the action of σ on X_2. Let X' be the normalization of the product of the $\bar{X}^\sigma$ over X_2. Then it is immediately seen that for this X', the covering V'^*/U'^* extends to an etale covering of X'^* above P.

In finding the map $X' \to X$ locally above a point P, it suffices to treat points above the closed point of Spec A. Since the closed fibres of X and X^* are the same, such a P has a single point of X^* lying above it, which we denote by the same letter. The problem being local for the etale topology, an easy limit argument shows that it suffices to prove the following: Let P be an arbitrary point of the closed fibre of $X/\operatorname{Spec} A$, let X_P (resp. X^*_P) denote the henselization of X (resp. X^*) at P. Then with the obvious notation, there is a proper surjective map $X' \to X_P$ such that the induced covering V'^* of $U'^* = X' \times_{X_P} U^*_P$ extends to an etale covering of X'^*.

Now it is easily seen that the local rings A_P (resp. A^*_P) of X_P (resp. X^*_P) have the same completion. Moreover the closed subscheme Y_P of X_P is the henselization of Y at P, and hence is isomorphic to Y^*_P. It is clear that if the theorem is true in this situation, then a map $X' \to X_P$ as above can be found. Thus we are reduced to proving the theorem in the case $X = \operatorname{Spec} A$.

I claim that we may moreover assume that

(*) The etale covering V^*/U^* is defined by restriction to U^* of $\operatorname{Spec} A^*[x]/(f)$ where $f(x) \in A^*[x]$ is a monic polynomial.

Suppose this is shown. Then the discriminant δ of f has the property that its zeros are contained in Y^*, since the extension is etale on $U^* = X^* - Y^*$. Thus the ideal (δ) contains J^{*n} for some n, J^* the ideal of Y^* in A. Now since $Y \approx Y^*$ and A, A^* have the same completion, the J^*-adic completion of A^* and the J-adic completion of A $(J = I(Y))$ are also isomorphic. Thus we can find an $f_1 \in A[x]$ such that in A^*

$$f_1 \equiv f \pmod{J^{*2n}\mathfrak{m}^*} \qquad (\mathfrak{m}^* = \max A^*)$$

hence

$$f_1 \equiv f \pmod{\delta^2\mathfrak{m}^*}.$$

By I. (1.3), the extensions $B^* = A^*[x]/(f)$ and $A^*[x]/(f_1)$ are isomorphic, hence B^* is induced from an extension B/A, which is what we want.

It remains to reduce ourselves to the case (*) above. What is clear is that locally at any $u^* \in U^*$ the extension is given by one equation ([5], I. 7. 5). We can take A normal. Then V^*/U^* is given by restriction from $\operatorname{Spec} B^*$ for some finite A^*-algebra B^*. Let $u \in U^*$, and say $b \in B^*_u$ generates the extension B^*_u/A^*_u. Then for some $r \in B$ invertible at u, $rb \in B^*$ and rb satisfies a monic equation of the correct degree over A^*. Therefore there exists a finite set $\{f_1, \cdots, f_n\}$ of monic polynomials in $A^*[x]$ such that for any $u \in U^*$ the extension B^*_u/A^*_u is isomorphic to $A^*_u[x]/(f_i)$ for at least one i.

Let δ_i be the discriminant of f_i, and $I^* = (\delta_1, \cdots, \delta_n)$. Then by construction it is clear that the support of $V(I^*)$ is contained in Y^*. Thus since the J-adic completion of A and the J^*-adic completion of A^* are isomorphic, the ideal I^* is of the form IA^* for some ideal I of A.

Now since we had just to prove the theorem for $X = \operatorname{Spec} A$, it is still permissible to make a proper surjective base change $X' \rightarrow X$, and then the problem is again local on X' for the etale topology. Let X' be the result of blowing up the ideal I in X. Then $X'^* = X' \times_X X^*$ is the result of blowing up I^*, and the ideal generated by I^* in X'^* is locally principal. Therefore at each point P' of X'^*, it is generated by one of the elements δ_i. But this implies that the extension defined by the equation f_i induces V'^*/U'^* in a neighborhood of P'. Since the problem is local on X' for the etale topology, this indeed reduces us to the case (*), and completes the proof.

3. Change of field for complete local rings. Let A be a complete local ring and $k \subset A$ a field such that $A/\mathfrak{m}A$ is finite and radicial over k. Let $L \supset k$ be an extension field and denote by A_L the completion of $A \otimes_k L$ with respect to the ideal generated by $\mathfrak{m}$. Denote by $\tilde{A}$ (resp. $\tilde{A}_L$) the strict localization, i.e., the maximal unramified extension, of A (resp. A_L).

PROPOSITION (3.1). *Let $\tilde{X} \subset \operatorname{Spec} \tilde{A}$ be an open subset, and denote by a subscript L the base change from $\tilde{A}$ to $\tilde{A}_L$. The functor $Et(\tilde{X}) \rightarrow Et(\tilde{X}_L)$ is fully faithful.*

Proof. Let A' be a finite A-module such that $f \colon \operatorname{Spec} A' \rightarrow \operatorname{Spec} A$ is surjective. Then f is a strict morphism of descent. Hence an argument similar to that used in the proof of (1.1) shows that we may replace A by A', hence that we may suppose A (and therefore $\tilde{A}$) normal. Let $\tilde{U}$, $\tilde{V}$ be etale coverings of $\tilde{X}$. They are obtained by restriction to $\tilde{X}$ from finite normal $\tilde{A}$-modules $\tilde{B}$, $\tilde{C}$, and the usual argument of the graph (cf. proof of (1.1)) reduces one to showing that $\operatorname{Spec} \tilde{B}$ is irreducible iff. $\operatorname{Spec} \tilde{B}_L$ is irreducible. But $\tilde{B}$ is normal, hence $\operatorname{Spec} \tilde{B}$ is geometrically unibranch, and so $\operatorname{Spec} \tilde{B}_L$ is also geometrically unibranch. For, replacing if necessary A by a finite unramified extension, $\tilde{B}$ is induced by a finite normal A-algebra B. Then B_L is geometrically unibranch by I. (2.5), hence $\tilde{B}_L$ is also geometrically unibranch (cf. (2.2)). Therefore it suffices to show that $\tilde{B}$ local implies $\tilde{B}_L$ local, which is trivial.

THEOREM (3.2). *With the above notation, suppose $\tilde{X}$ is obtained from $\operatorname{Spec} \tilde{A}$ by removing the closed point and that every irreducible component of $\operatorname{Spec} A$ is of dimension $\geqq 2$. Then the functor $Et(\tilde{X}) \rightarrow Et(\tilde{X}_L)$ is an equivalence of categories.*

Proof. If $L \supset K \supset k$, and if the theorem is true for the extensions L/K and K/k, it is clearly true for L/k. Moreover, the case of a finite extension

K/k is trivial—for a separable extension the strict localizations are equal, and for an inseparable extension the map $\bar{X}_L \rightarrow \bar{X}$ is radicial. Thus to prove the theorem for any extension L/k of finite type, it suffices to treat the case that $L = k(t)$ is a pure transcendental extension. By descent and (3.1), it suffices to treat the case that A is normal.

Assume that $L = k(t)$. Now $\tilde{A}_L = \varinjlim A_{L'}$ where L' runs through finite separable extensions of L. Hence if $\bar{U}_L/\bar{X}_L$ is an etale covering, it is induced from an etale covering of $\operatorname{Spec} A_{L'} - \{a\}$ (a the closed point) for some such L'. Now the local ring A' of $A \otimes_k L'$ at the ideal generated by $\mathfrak{m} = \max A$ is localized from an algebra of finite type over A, and hence is an excellent local ring ([4], IV. 7. 8. 6 (i)). Therefore by (2.1) the etale covering $\bar{U}_L$ is induced by and etale covering of $\operatorname{Spec} A_{L'}{}^* - \{a\}$ where $A_{L'}{}^*$ is the henselization of $A \otimes_k L'$ at the ideal generated by $\mathfrak{m}$. This henselization is a limit of rings etale and of finite type over $A \otimes_k k[t] = A[t]$ (since L' is separable over $L = k(t)$). Therefore the covering is induced from some such ring B by an etale covering of $\operatorname{Spec} B - \operatorname{Spec}(B/\mathfrak{m}B)$. Replacing if necessary A by $A_{k'}$ for a finite separable extension k' of k, we may assume that $\operatorname{Spec}(B/\mathfrak{m}B)$ contain a point P rational over $A/\mathfrak{m}$. Let B^* denote the henselization of B at P. This ring is isomorphic to the henselization $A\{t\}$ of $A[t]$ at the "origin," and we have a canonical map $B^* \rightarrow A_{L'}{}^*$. Our given covering is induced from an etale covering U^* of

$$X^* = \operatorname{Spec} B^* - \operatorname{Spec}(B^*/\mathfrak{m}B^*),$$

and we have to show that U^* is induced by an etale covering U of X. But this is implied by ([1], XV (3.2)). For, choose an etale covering $\bar{U}$ of $\bar{X}$ such that $U^* \times_{\bar{X}} \bar{V}$ is completely decomposed over $V^* = X \times_{\bar{X}} \bar{V}$ on the section $\{t = 0\}$. $\bar{V}$ is induced by a finite normal A-algebra A', and $V^* = \operatorname{Spec} A'\{t\} - \operatorname{Spec} A'\{t\}/\mathfrak{m}$. Thus ([1], XV (3.2)) implies that $U^* \times_{\bar{X}} \bar{V}$ is completely decomposed over V^*, which gives the result. This proves the theorem for an extension L/k of finite type.

Consider now the general case. By descent and (3.1), we may as above assume A normal, hence that $\operatorname{Spec} A$ is geometrically unibranch. Moreover, applying I. (2.5) and the above result, we may replace A by A_K where K is any subextension of L of finite type over k. Let $\bar{U}_L/\bar{X}_L$ be an etale covering, say galois. Replacing if neecssary L by a finite separable extension L', we may suppose that $\hat{U}_L$ is induced by an etale galois covering U_L of $X_L = \operatorname{Spec} A_L - \{a\}$.

We use an argument similar to that of (2.1). Let C_L be a finite A_L-algebra which induces the etale covering U_L of X_L, and let $c_1, \cdots, c_n$ be elements of C_L such that locally at any point of X_L one of them generates the etale covering U_L, and such that c_i satisfies a monic equation $f_i(x) \in A_L[x]$ of the correct degree (cf. proof of (2.1)). Let δ_i be the discriminant of $f_i(x)$. The ideal $J_L = (\delta_1 \cdot \cdot \cdot, \delta_n)$ is $\mathfrak{m}_L$-primary ($\mathfrak{m}_L = \mathfrak{m} A_L$), i.e., is a quotient of $A_L/\mathfrak{m}_L{}^n$ for large n. It is easily seen that therefore J_L is induced from an ideal in A_K for some K/k of finite type. We may assume $K = k$.

Let $Y \rightarrow \operatorname{Spec} A$ be the result of blowing up the ideal $J = J_k$. Replacing k by an extension of finite type, we may assume that there are polynomials $f_i{}^0(x) \in A[x]$ such that at any point P on the closed fibre of Y_L,

$$f_i{}^0(x) \equiv f_i(x) \qquad (\operatorname{mod} \delta_i{}^2 I(P))$$

for some i for which δ_i defines the (locally principal) ideal generated by J_L at P. For, to do this it suffices to find $f_i{}^0$ congruent f_i modulo $J^2\mathfrak{m}$ in A, which is clearly possible after a finitely generated extension. Then in the henselization B of Y_L at P, the extension defined by the equations f_i and $f_i{}^0$ are isomorphic (I.(1.3)). Therefore if A_{Li} is the extension of A_L defined by $f_i{}^0$, the extension defined by f_i induces a completely decomposed extension of $B_i = B \otimes_{A_L} A_{Li}$ outside of the locus $\{\delta_i = 0\}$. Let A' be the normalization of $A[\xi_1, \cdot \cdot \cdot, \xi_n]$ where ξ_i satisfies the equation $f_i{}^0(\xi_i) = 0$, and let Y' be the normalization of Y in the ring of rational functions $R(A')$. Let $C'_L = C_L \otimes_A A'$ (where C_L is the A_L-algebra defined above). Then above any point P on the closed fibre of Y_L, the extension defined by C'_L over Y'_L is completely decomposed outside of the closed fibre, in the henselization at P. Therefore the normalization of $Y'_L \times_{\operatorname{Spec} A'_L} \operatorname{Spec} C'_L = \bar{Z}'_L$ is etale over the normalization $\bar{Y}'_L$ of Y'_L. But by I.(2.5) Y'_L is geometrically unibranch and so $\bar{Y}'_L \rightarrow Y'_L$ is radicial. Therefore there exists a finite etale covering Z'_L of Y'_L which is isomorphic to that induced by C'_L outside of the closed fibre. Now the category of etale coverings of Y'_L and of its closed fibre are equivalent by ([1], XII(5.9)). Moreover the category of etale coverings of a proper scheme over a field is invariant under separably closed ground field change ([5], X.1.8). Therefore, replacing if necessary k by a finite extension k', the covering Z'_L is induced by an etale covering of Y', and so the etale covering defined by C'_L on X'_L is induced by an etale covering of X'. Thus we have induced the original covering $\hat{U}_L/\hat{X}_L$ after a finite extension $\tilde{A} \rightarrow \tilde{A}'$ such that $f \colon \operatorname{Spec} \tilde{A}' \rightarrow \operatorname{Spec} \tilde{A}$ is surjective. By (3.1) we also obtain descent data for the map f, and so the theorem follows from ([5], IX).

4. The equal characteristic case.

LEMMA (4.1). *Let $A \rightarrow B$ be a morphism of excellent equicharacteristic hensel rings which is regular* ([4], IV.6.8.1). *Suppose* Spec A *geometrically unibranch and of dimension* $\geqq 2$. *Let $\tilde{A}$ (resp. $\tilde{B}$) be the strict localizations* *of A (resp. B). Then every etale covering of* $\operatorname{Spec}\tilde{B}-\{b\}$ *is induced from an etale covering of* $\operatorname{Spec}\tilde{A}-\{a\}$ *(a and b denote the closed points).*

Proof. We begin with some preliminary constructions. Since the normalization of A is a radicial extension ([4], 0_{IV}. 23.2.2), we may assume A (therefore B ([4], IV. 6.8.3(iii)) is normal. In characteristic zero, it is easy to see that A and B have compatible fields of representatives. Suppose that the characteristic is p different from zero. Choose a p-basis for $k=A/\mathfrak{m}$ and lift it to A. This p-basis determines a coefficient field in $\hat{A}$. Now it is clear that $\hat{B}/\mathfrak{m}\hat{B}$ is regular over k. Therefore I.(3.1) applies. Let k'/k be the radicial extension of this proposition. Since the p-basis for k was chosen in A, we may assume that there is a radicial extension A' of A such that $\hat{A}'\approx\hat{A}\otimes_k k'$. In characteristic zero, set $A=A'$. Then there are elements $z_1,\cdots,z_d$ of $\hat{B}'=A'\otimes_A\hat{B}$ which reduce to a regular system of parameters of $\hat{B}'/\hat{\mathfrak{m}}'\hat{B}'$ and such that a coefficient field $k'\subset\hat{A}'$ extends to a coefficient field $L'\subset\hat{B}'/(z_1,\cdots,z_d)$.

Consider the completion $\hat{A}'_{L'}$ of $\hat{A}'\otimes_k L'$ at the ideal generated by $\hat{\mathfrak{m}}'=\max\hat{A}'$. This ring is geometrically unibranch by I.(2.5), and there is a canonical map $\phi\colon\hat{A}'_{L'}\to\hat{B}'/(z_1,\cdots,z_d)$. These two complete local rings have the same residue fields. Moreover $\hat{\mathfrak{m}}'$ generates the maximal ideal of both rings. Therefore $\hat{B}'/(z_1,\cdots,z_d)$ as a quotient of $\hat{A}'_{L'}$. But $\hat{A}'_{L'}$ is of the same dimension as $\hat{B}'/(z_1,\cdots,z_d)$, and is irreducible. Thus the kernel of ϕ is nilpotent. This implies that the two rings have the same etale extensions. Now it is easily seen that nothing essential is altered in our construction if we replace A and B by finite etale extensions A_1, B_1 with a compatible map. We leave the verification to the reader. Hence it is still permissible to make such a change.

We can now complete the proof. Let $\hat{U}$ be an etale covering of $\operatorname{Spec}\hat{B}-\{b\}$. Replacing if necesssary B by a finite etale extension, we may suppose $\hat{U}$ induced by an etale covering U of $\operatorname{Spec}B-\{b\}$. This covering determines an etale covering $\hat{U}$ of $\operatorname{Spec}\hat{B}'-\{b\}$, hence a covering $\hat{U}'$ of $\operatorname{Spec}\hat{A}'_{L'}-\{b\}$ via the above construction. If we apply (3.2) to the ring $\hat{A}'$ and the field extension $k'\subset L'$, it follows that, replacing if necessary A (resp. B) by finite etale extension, we may assume that $\hat{U}'$ is induced from an etale covering $\hat{V}'$ of $\operatorname{Spec}\hat{A}'-\{a'\}$. By (2.1), $\hat{V}'$ is induced from an etale covering V' of $\operatorname{Spec}A'-\{a'\}$, hence by an etale covering V of $\operatorname{Spec}A-\{a\}$. Let $X=\operatorname{Spec}B-\operatorname{Spec}B/\mathfrak{m}B$. It suffices to show that the etale covering of X induced by V and that induced by U are isomorphic. By (1.1) and the fact that B' is a radicial extension of B, it suffices to show that the coverings induced on $\hat{X}'=\operatorname{Spec}\hat{B}'-\operatorname{Spec}(\hat{B}'/\mathfrak{m}'\hat{B}')$ are isomorphic, and we know by construction that the induced coverings of $\operatorname{Spec}\hat{B}'/(z_1,\cdots,z_d) - \{b'\}$ are isomorphic. Put $\hat{B}'_i = \hat{B}'/(z_1,\cdots,z_i)$ and $Z_i = \operatorname{Spec}\hat{B}'_i - \{b'\}$ $(i = 1,\cdots,d)$. By SGA '62 Exp. X Cor. 2.5, it suffices to show that, in the terminology of SGA '62, $\operatorname{Lef}(Z_1, Z_d)$ holds. We will show inductively that $\operatorname{Lef}(Z_i, Z_{i+1})$ holds for $i = 1,\cdots,d$. By SGA '62 X Example 2.1, it suffices to verify the following two conditions:

(a) $z_1,\cdots,z_d$ is a regular sequence in $\hat{B}'$.

(b) For each point $p' \in \operatorname{Spec}\hat{B}'_i$ with $\dim \hat{B}'/p' = 1$, and each $i < d$, $\operatorname{depth}_{p'}(\hat{B}'_i) \geqq 2$.

Now it is easily seen that $\hat{B}'$ is flat over $\hat{A}'$ (SGA '60 IV.5.9). Since $z_1,\cdots,z_d$ is a regular sequence $(\operatorname{mod} \mathfrak{m}'\hat{B}')$, it is a regular sequence in $\hat{B}'$ by (SGA '60 IV.5.7). Thus (a) holds. To show (b), first suppose the image p'_0 of p' in $\operatorname{Spec}\hat{A}'$ has codimension $\geqq 2$. Then since A is normal (hence $\hat{A}$ is normal) and $\hat{A}'$ is flat over A, $\operatorname{depth}_{p'_0}\hat{A}' \geqq 2$. Since because of (a) $\hat{B}'_i$ is flat over $\hat{A}'$, $\operatorname{depth}_{p'}\hat{B}'_i \geqq 2$. If the image p'_0 is of codimension $\leqq 1$, then p'_0 lies over a regular point of $\operatorname{Spec}\hat{A}$. Therefore the image in $\operatorname{Spec} A$ is regular, hence its inverse image in $\operatorname{Spec}\hat{B}$ is regular. Thus p' lies over a regular point p of $\operatorname{Spec}\hat{B}$. Therefore

$$\operatorname{depth}_{p'}\hat{B}' = \dim \hat{B}' - 1 = \dim A + d - 1.$$

Since $(z_1,\cdots,z_d)$ is an $\hat{\mathfrak{m}}'$-regular sequence, it is also a p'-regular sequence. Thus

$$\operatorname{depth}_{p'}\hat{B}'_i = \dim A + d - 1 - i \geqq \dim A \geqq 2$$

as required. This completes the proof.

THEOREM (4.2). *Let A be an excellent equicharacteristic hensel ring, $X/\operatorname{Spec} A$ a proper scheme, $Y \subset X$ a closed subscheme, and $U = X - Y$. Suppose that in the normalization $\bar{X}$ of X every point of $\bar{Y} = Y \times_X \bar{X}$ has codimension $\geqq 2$. Then $Et(U) \to Et(\hat{U})$ is an equivalence of categories.*

Proof. The functor is fullly faithful by (1.1). Let $\hat{V}/\hat{U}$ be an etale covering. To show that $\hat{V}$ is induced from an etale covering V of U, it suffices by (1.1) and descent ([5], IX (4.12)) to find $X' \to X$ proper such that $\hat{V}' = \hat{V} \times_X X'$ is induced from an etale covering V' of $U' = U \times_X X'$. Since X' and $\hat{X}'$ have the same closed fibre, it suffices by ([1], XII (5.9)) to find X' such that $\hat{V}'$ extends to an etale covering of $\hat{X}'$. As in the proof of (2.1), one sees that therefore the problem is local on X for the etale topology.

We may assume X normal. Let $\hat{y} \in \hat{Y}$ be a point of maximal dimension, say d, at which the extension defined by $\hat{V}/\hat{U}$ ramifies, and let y be the image of $\hat{y}$ in Y. Let B be the henselization of X at y and C the henselization of $\hat{X}$ at $\hat{y}$. Since A is excellent, the map $\hat{X} \to X$ is regular ([4], IV. 7. 8. 1) and it is easy to see that therefore the map $B \to C$ is also a regular map, and B, C are excellent rings. Therefore by (4. 1) the etale covering of $\operatorname{Spec} C - \{c\}$ defined by $\hat{V}$ is induced by an extension B' of B. This implies (cf. proof of (2. 1)) that there exists a finite surjective map $X' \to X$ with X' normal such that the extension defined by $\hat{V}$ on $\hat{U}'$ is unramified above any point of dimension d on $\hat{Y}$, hence is unramified at any point of $\hat{Y}'$ of dimension $\geqq d$. By induction on d, there exists $X' \to X$ finite and surjective such that the etale covering $\hat{V}'$ of $\hat{U}'$ extends to an etale covering of $\hat{X}'$. This completes the proof.

5. Variants. Let A be an excellent equicharacteristic hensel ring, $X/\operatorname{Spec} A$ proper, $Y \subset X$ a closed subscheme, and $U = X - Y$. Let $\hat{V}$ be an etale covering of $\hat{U}$. In order that $\hat{V}$ be induced by an etale covering V of U, it suffices by descent and (4. 2) that there exist an $X' \to X$ proper surjective with X' normal so that the covering $\hat{V}'$ is unramified above the points of $\hat{Y}'$ of codimension 1. Applying Abyankhar's lemma, we get in particular the following result:

THEOREM (5. 1). *With the above notation, assume X normal. If the covering $\hat{V}$ is tamely ramified at points of $\hat{Y}$ which are of codimension 1 in $\hat{X}$, then $\hat{V}$ is induced by an etale covering V of U.*

Since every extension is tamely ramified in characteristic zero, this gives an alternative proof of (1. 3) making no use of resolution of singularities.

Another variant is

THEOREM (5. 2). *Let A be an excellent equicharacteristic hensel ring and $X/\operatorname{Spec} A$ a proper normal irreducible scheme. Let $\hat{f}\colon \hat{X}' \to \hat{X}$ be a finite morphism with $\hat{X}'$ normal, and suppose $\hat{f}$ is etale at points of $\hat{X}'$ of codimension 1. Then $\hat{f}$ is induced by a morphism $f\colon X' \to X$.*

Proof. By purity ([5], X. 3. 6), $\hat{f}$ is etale at every regular point of $\hat{X}$. Let $Y \subset X$ be its singular locus and $U = X - Y$. Since X is normal, Y is everywhere of codimension $\geqq 2$ in X. Therefore (4. 2) applies and the restriction $\hat{V}$ of $\hat{X}'$ to $\hat{U}$ is induced by an etale covering V of U. Let X' be the normalization of X in $R(V)$. Then $X' \times_X \hat{X}$ is normal ([4], IV. 6. 8. 3), hence $X' \times_X \hat{X} \approx \hat{X}'$.

If we specialize the results (1. 3), (2. 1), (4. 2) to the case $X = \operatorname{Spec} A$, we obtain

COROLLARY (5. 3). *Let A be an excellent hensel ring and U an open* *subset of* Spec$A = X$. *Then* $Et(\hat{U}) \rightarrow Et(\hat{U})$ *is an equivalence of categories in each of the following cases:*

(i) *The residue characteristic of* A *is zero.*

(ii) A *is an equicharacteristic ring, and every point of* $Y = X - U$ *outside of the closed point has codimension at least* 2 *in each irreducible component of* X *passing through it.*

(iii) U *is the complement of the closed point in* X.

Theorem (5.2) for the case $X = \operatorname{Spec} A$ reads:

COROLLARY (5.4). *Let* A *be an excellent equicharacteristic normal hensel ring, and let* $\hat{B}$ *be a finite normal* $\hat{A}$*-algebra which is etale over* $\hat{A}$ *in codimension* 1. *Then* $\hat{B} = B \otimes_A \hat{A}$ *for some finite* A*-algebra* B.

Using (4.2) and the method of proof of (4.1), we obtain

COROLLARY (5.5). *Let* $A \rightarrow B$ *be a regular morphism of excellent equicharacteristic hensel rings. Let* U_A *be the complement of the closed point of* Spec$A$, *and* U_B *the inverse image of* U_A *in* Spec$B$. *Then* $Et(U_A) \rightarrow Et(U_B)$ *is an equivalence of categories.*

We leave the verification of this corollary to the reader.

REFERENCES.

[1] M. Artin and A. Grothendieck. *Séminaire de Géométrie Algébrique* 1963-4. *Cohomologie Étale des Schemas,* Institut des Hautes Études Scientifiques (mimeographed notes).

[2] N. Bourbaki, *Algèbre Commutative,* Chapitre III. Paris (1961).

[3] I. S. Cohen, "On the structure and ideal theory of complete local rings," *Transactions of the American Mathematical Society,* vol. 59, No. 1 (1946), pp. 54-106.

[4] A. Grothendieck and J. Dieudonné, *Éléments de Géométrie Algébrique,* Publications Mathématiques de l'Institut des Hautes Études Scientifiques (1960-).

[5] A. Grothendieck, *Séminaire de Géométrie Algébrique* (1960-1), Institut des Hautes Études Scientifiques (mimeographed notes).

[6] ———, *Séminaire de Géométrie Algébrique* (1962), Institut des Hautes Études Scientifiques (mimeographed notes).

[7] R. Hartshorne, "Complete intersections and connectedness," *American Journal of Mathematics,* vol. 84, No. 3 (1962). pp. 479-508.

[8] H. Hironaka, "Resolution of singularities of an algebraic variety over a field of characteristic zero," I, II, *Annals of Mathematics,* vol. 79 Nos. 1, 2 (1964).

[9] M. Nagata, *Local Rings,* New York (1962).

[10] M. Artin, "On algebraic extensions of local rings," *Rendiconti di Matematica* (to appear).

ON A PROBLEM OF CALKIN.

By Shôichirô Sakai.*

---

1. Introduction. Let $\mathfrak{H}$ be a separable Hilbert space, $B(\mathfrak{H})$ the C^*-algebra of all bounded operators on $\mathfrak{H}$, $C(\mathfrak{H})$ the C^*-algebra of all compact operators on $\mathfrak{H}$. Calkin [1] studied the *-representations of the quotient C^*-algebra $B(\mathfrak{H})/C(\mathfrak{H})$, and raised the problem of whether or not there exists a type III-factor *-representation of $B(\mathfrak{H})/C(\mathfrak{H})$.

In the present paper, we shall show that this algebra has a type III-factor *-representation. The crucial tools in the proof of the theorem are the results of the author [5] and Schwartz [7].

It is still open whether or not the C^*-algebra has a type II_∞-factor *-representation (even a global type II_∞ *-representation).

2. The theorem. In this section, we shall show the following theorem.

Theorem. *The C^*-algebra $B(\mathfrak{H})/C(\mathfrak{H})$ has a type* III-*factor *-representation.*

To prove the theorem, we shall provide some considerations. Schwartz [7, 8] showed that there is a type III-factor M in the Hilbert space $\mathfrak{H}$ satisfying the following: there is a linear projection P with norm 1 of $B(\mathfrak{H})$ onto M satisfying the conditions:

(1) $P(a^*) = P(a)^*$ for $a \in B(\mathfrak{H})$;

(2) $P(h) \geqq 0$ if $h \geqq 0$ and $h \in B(\mathfrak{H})$;

(3) $P(axb) = aP(x)b$ for $a, b \in M$ and $x \in B(\mathfrak{H})$.

The projection P also satisfies the following condition:

(4) $P(x^*x) \geqq P(x)^*P(x)$ for $x \in B(\mathfrak{H})$.

Now let Ω be the set of all linear projections of $B(\mathfrak{H})$ onto M satisfying the conditions (1), (2), (3) and (4). Then every element Q of Ω is a bounded linear mapping of $B(\mathfrak{H})$ onto M and its norm on the self-adjoint portion of $B(\mathfrak{H})$ is one—in fact, let k be a self-adjoint element of $B(\mathfrak{H})$,

---

Received November 4, 1965.

* This paper written with partial support from ONR contract ONR NR-551 (57).

then $\| k \| \cdot I \pm k \geqq 0$ and so $\Omega(\| k \| \cdot I \pm k) = \| k \| \cdot I \pm Q(k) \geqq 0$, where I is the identity operator on $\mathfrak{H}$; hence $\| Q(k) \| \leqq \| k \|$.

Let $\mathfrak{L}(B(\mathfrak{H}), M)$ be the Banach space of all bounded linear mappings of $B(\mathfrak{H})$ into M. Then it is the dual of the Banach space $B(\mathfrak{H}) \otimes_{\gamma} M_{*}$ (cf. (3)), where M_{*} is the associated space of M (namely, the dual of $M_{*} = M$) and γ is the greatest cross norm.

Lemma 1. *Ω is a $\sigma(\mathfrak{L}(B(\mathfrak{H}), M), B(\mathfrak{H}) \otimes_{\gamma} M_{*})$—compact convex subset of $\mathfrak{L}(B(\mathfrak{H}), M)$.*

Proof. It is clear that all elements of the clousure $\bar{\Omega}$ of Ω are linear projections satisfying the conditions (1), (2), and (3).

Let $Q \in \bar{\Omega}$, then $\langle Q(x)^* Q(x), \phi \rangle = \langle Q(x^* Q(x)), \phi \rangle = \langle x^* Q(x), Q^*(\phi) \rangle$ for $x \in M$ and $\phi \ (\geqq 0) \in M_{*}$, where Q^* is the dual of Q and $\langle (\cdot), \phi \rangle$ is the value of ϕ at the point $(\cdot)$. $Q^*(\phi)$ is a positive linear functional on $B(\mathfrak{H})$, because $\langle h, Q^*(\phi) \rangle = \langle Q(h), \phi \rangle \geqq 0$ for all positive $h \in B(\mathfrak{H})$. Hence,

$$\langle x^* Q(x), Q^*(\phi) \rangle \leqq \langle x^* x, Q^*(\phi) \rangle^{\frac{1}{2}} \langle Q(x)^* Q(x), Q^*(\phi) \rangle^{\frac{1}{2}}.$$

Therefore $\langle Q(x)^* Q(x), \phi \rangle \leqq \langle Q(x^* x), \phi \rangle$, because

$$\langle Q(x)^* Q(x), Q^*(\phi) \rangle = \langle Q(Q(x)^* Q(x)), \phi \rangle = \langle Q(x)^* Q(x), \phi \rangle;$$

hence $Q(x)^* Q(x) \leqq Q(x^* x)$. The convexity of Ω is clear. This completes the proof.

Let ϕ be a normal, faithful, positive linear functional on M (namely, $\phi(a^* a) = 0$ implies $a = 0$ for $a \in M$ such that $\phi(I) = 1$. For $Q \in \Omega$, we shall define a positive linear functional ϕ_Q on $B(\mathfrak{H})$ by $\phi_Q(x) = \phi(Q(x))$ for $x \in B(\mathfrak{H})$.

Let $\mathcal{E} = \{\phi_Q \mid Q \in \Omega\}$, then we can easily see that $\mathcal{E}$ is a compact convex subset of the state space of $B(\mathfrak{H})$ with the topology $\sigma(B(\mathfrak{H})^*, B(\mathfrak{H}))$, where $B(\mathfrak{H})^*$ is the dual of $B(\mathfrak{H})$.

Lemma 2. *$Q(C(\mathfrak{H})) = 0$ for all $Q \in \Omega$.*

Proof. Suppose that there are Q_0 in Ω and h_0 in $C(\mathfrak{H})$ such that $Q_0(h_0) \neq 0$. It is enough to assume that $h_0 > 0$. Then, $Q_0(h_0) > 0$; therefore we can choose a non-zero projection e in M such that $\lambda e \leqq Q_0(h_0)$ for some positive number λ and e commutes with $Q_0(h_0)$. Then $Q_0(Meh_0) = MeQ_0(h_0) \supset Me$; therefore $Q_0(Meh_0)$ is not uniformly separable. On the other hand, $Meh_0 \subset C(\mathfrak{H})$ and so it is uniformly separable and it is a contradiction. This completes the proof.

Let $\phi_Q (Q \in \Omega)$ be an extrem point of $\mathcal{E}$, and let $\{\pi_Q, \mathfrak{H}_Q\}$ be the

*-representation of $B(\mathfrak{H})$ on a Hilbert space $\mathfrak{H}_Q$ constructed via ϕ_Q. By Lemma 2, $\phi_Q(C(\mathfrak{H}))=0$ and so $\{\pi_Q, \mathfrak{H}_Q\}$ can be canonically considered a *-representation of $B(\mathfrak{H})/C(\mathfrak{H})$.

Let N be the weak closure of $\pi_Q(B(\mathfrak{H}_Q))$ on $\mathfrak{H}_Q$. Then we shall define a linear mapping of N onto M in the following consideration.

Let N_* be the associated space of N. For $f \in M_*$, we define $F(\pi_Q(x)) = f(Q(x))$ for $x \in B(\mathfrak{H})$; then F is a bounded linear functional on the C^*-algebra $\pi_Q(B(\mathfrak{H}))$.

First of all, we shall show that F is strongly continuous on bounded spheres. Let $\{\pi_Q(x_\alpha) \mid x_\alpha \in B(\mathfrak{H}), \| x_\alpha \| \leqq 1\}$ be a directed set of elements in $\pi_Q(B(\mathfrak{H}))$ which converges strongly to 0, then $\phi_Q(x_\alpha{}^* x_\alpha) = \phi(Q(x_\alpha{}^* x_\alpha)) \Rightarrow 0$; hence $\phi(Q(x_\alpha{}^* x_\alpha)) \geqq \phi(Q(x_\alpha)^* Q(x)_\alpha) \Rightarrow 0$. Therefore, $\{Q(x_\alpha)\}$ converges strongly to 0 in the W^*-algebra M, because $\{Q(x_\alpha)\}$ is bounded and ϕ is normal, faithful; hence

$$f(Q(x_\alpha)) = F(\pi_Q(x_\alpha)) \Rightarrow 0.$$

LEMMA 3. *Let G be a bounded linear functional of $\pi_Q(B(\mathfrak{H}))$ which is strongly continuous on bounded spheres, then it can be uniquely extended to a normal linear functional $\bar{G}$ on N with $\| G \| = \| \bar{G} \|$.*

Proof. Let $\pi_Q(B(\mathfrak{H}))^{**}$ be the second dual of the C^*-algebra $\pi_Q(B(\mathfrak{H}))$, then it is a W^*-algebra (cf. [9]). For $H \in N_*$, let $\tilde{H}$ be the restriction of H on $\pi_Q(B(\mathfrak{H}))$, then $\| \tilde{H} \| = \| H \|$ by the density theorem of Kaplansky (cf. [2]). Put $\tilde{N}_* = \{\tilde{H} \mid H \in N_*\}$, then $\tilde{N}_*$ is a closed subspace of the dual $\pi_Q(B(\mathfrak{H}))^*$ of $\pi_Q(B(\mathfrak{H}))$ and moreover it is invariant—namely $\tilde{H} \in \tilde{N}_*$ implies $L_a R_b \tilde{H} \in \tilde{N}_*$ for $a, b \in \pi_Q(B(\mathfrak{H}))$, where $L_a R_b \tilde{H}(x) = \tilde{H}(axb)$ for $x \in \pi_Q(B(\mathfrak{H}))$. Let $\tilde{N}_*{}^0$ be the polar of $\tilde{N}_*$ in $\pi_Q(B(\mathfrak{H}))^{**}$, then it is a σ-weakly closed ideal (cf. [4]); therefore, there is a central projection z of $\pi_Q(B(\mathfrak{H}))^{**}$ such that $\tilde{N}_*{}^0 = \pi_Q(B(\mathfrak{H}))^{**} z$. Moreover we can easily see that the W^*-algebra N (resp. the associated space N_*) can be canonically identified with $\pi_Q(B(\mathfrak{H}))^{**}(I-z)$ (resp. $\tilde{N}_*$).

Suppose that Lemma 3 is not true. Then there is no element $\bar{G}$ in N_* satisfying $G = \bar{G}$ on $\pi_Q(B(\mathfrak{H}))$; therefore $L_z G \neq 0$; hence there is an element a in $\pi_Q(B(\mathfrak{H}))^{**}$ such that $az = a$ and $G(a) \neq 0$.

Now, choose a bounded directed set (a_α) of elements in $\pi_Q(B(\mathfrak{H}))$ such that $a_\alpha \to a$ (strongly) in the W^*-algebra $\pi_Q(B(\mathfrak{H}))^{**}$, then $a_\alpha(I-z)$ converges strongly to $a(I-z) = 0$; hence by the assumption, $G(a_\alpha) \to 0$. On the other hand, $G(a_\alpha) \to G(a) \neq 0$. This is a contradiction. This completes the proof.

By Lemma 3, the F can be uniquely extended to an element $\bar{F}$ of N_* with $\| F \| = \| \bar{F} \|$. Put $T(f) = \bar{F}$ for $f \in M_*$, then T is a bounded linear mapping of M_* into N_*. Let T^* be the dual of T, then T^* is a continuous linear mapping of N with the topology $\sigma(N, N_*)$ into M with the topology $\sigma(M, M_*)$.

Lemma 4. *T^* satisfies the following conditions:*

(1) $T^*(\pi_Q(x)) = Q(x)$ *for* $x \in B(\mathfrak{H})$,

(2) $T^*(y^*) = T^*(y)^*$ *for* $y \in N$,

(3) $T^*(h) \geqq 0$ *for* $h\ (\geqq 0) \in N$,

(4) $T^*(\pi_Q(a) y \pi_Q(b)) = aT^*(y)b$ *for* $a, b \in M$ *and* $y \in N$,

(5) $T^*(y^*y) \geqq T^*(y)^* T^*(y)$ *for* $y \in N$.

Proof.

(1) $\langle T^*(\pi_Q(x)), f\rangle = \langle \pi_Q(x), T(f)\rangle = \langle Q(x), f\rangle$ for $f \in M_*$. Hence, $T^*(\pi_Q(x)) = Q(x)$.

(2) Let (x_α) be a bounded directed set of elements of $B(\mathfrak{H})$ such that $\pi_Q(x_\alpha) \to y$ (σ-weakly) in N, then

$$T^*(\pi_Q(x_\alpha)^*) = T^*(\pi_Q(x_\alpha^*)) = Q(x_\alpha^*) = Q(x_\alpha)^* = T^*(\pi_Q(x_\alpha))^* \to T^*(y)^*$$

(σ-weakly) in M, and $T^*(\pi_Q(x_\alpha)^*) \to T^*(y^*)$ (σ-weakly) in M; hence $T^*(y)^* = T^*(y^*)$.

(3) The proof is analogous with (2).

(4) $T^*(\pi_Q(a)\pi_Q(x)\pi_Q(b)) = T^*(\pi_Q(axb)) = Q(axb) = aQ(x)b$ for a, $b \in M$ and $x \in B(\mathfrak{H})$. Therefore, by the σ-weak continuity of T^* and the density of $\pi_Q(B(\mathfrak{H}))$ in N, we can prove this easily.

(5) Let χ be a positive element of M_*, then

$$\begin{aligned}\langle T^*(y)^* T^*(y), \chi\rangle &= \langle T^*(y^*\pi_Q(T^*(y))), \chi\rangle \\ &= \langle y^*\pi_Q(T^*(y)), T(\chi)\rangle \leqq \langle y^*y, T(\chi)\rangle^{\frac{1}{2}} \langle \pi_Q(T^*(y)^*)\pi_Q(T^*(y)), T(\chi)\rangle^{\frac{1}{2}} \\ &= \langle y^*y, T(\chi)\rangle^{\frac{1}{2}} \langle T^*(y)^*T^*(y), \chi\rangle^{\frac{1}{2}}\end{aligned}$$

because $\langle \pi_Q(h), T(\chi)\rangle = \langle Q(h), \chi\rangle$ for $h \in B(\mathfrak{H})$ and so $T(\chi) \geqq 0$; hence $T^*(y)^*T^*(y) \leqq T^*(y^*y)$. This completes the proof.

Lemma 5. *N is a factor.*

Proof. Let Z_p be the set of all central projections of N. For $z \in Z_p$, $T^*(z)$ is a central element of M, because $T^*(z)a = T^*(z\pi_Q(a)) = T^*(\pi_Q(a)z)$

$= aT^*(z)$ for $a \in M$. Therefore, $T^*(z) = \lambda(z)I$, where $\lambda(z)$ is a real number such that $0 \leqq \lambda(z) \leqq 1$.

Suppose that $0 < \lambda(z) < 1$, then we shall define two linear mappings Q_1 and Q_2 of $B(\mathfrak{H})$ into M as follows: $Q_1(x) = (1/\lambda(z)) \cdot T^*(\pi_Q(x)z)$ and $Q_2(x) = (1/1 - \lambda(z)) \cdot T^*(\pi_Q(x)(I - z))$ for $x \in B(\mathfrak{H})$. Then Q_1 (resp. Q_2) is a projection of $B(\mathfrak{H})$ onto M, because $Q_1(a) = (1/\lambda(z))T^*(\pi_Q(a)z) = (1/\lambda(z))aT^*(z) = a$ for $a \in M$ (resp. $Q_2(a) = a$).

Moreover,

$$Q_1(x)^* = (1/\lambda(z))T^*(\pi_Q(x)z)^* = (1/\lambda(z))T^*(\pi_Q(x^*)z) = Q_1(x^*)$$

for $x \in B(\mathfrak{H})$; $Q_1(h) = (1/\lambda(z))T^*(\pi_Q(h)z) \geqq 0$ for h $(\geqq 0) \in B(\mathfrak{H})$;

$$Q_1(axb) = (1/\lambda(z))T^*(\pi_Q(axb)z) = (1/\lambda(z)aT^*(\pi_Q(x)z)b = aQ_1(x)b$$

for $a, b \in M$ and $x \in B(\mathfrak{H})$. Therefore Q_1 and analogously Q_2 belong to Ω (cf. the proof of Lemma 1). Then,

$$\begin{aligned} \langle x, \phi_{Q_1} \rangle &= \langle (1/\lambda(z))T^*(\pi_Q(x)z), \phi \rangle \\ &= (1/\lambda(z)) \langle \pi_Q(x)z, T(\phi) \rangle. \end{aligned}$$

Analogously, $\langle x, \phi_{Q_2} \rangle = (1/1 - \lambda(z)) \langle \pi_Q(x)(I - z), T(\phi) \rangle$. Therefore, $\phi_Q = \lambda(z)\phi_1 + (1 - \lambda(z))\phi_{Q_2}$; by the extremity of ϕ_Q, $\phi_Q = \phi_{Q_1} = \phi_{Q_2}$ on $B(\mathfrak{H})$ and so $1/\lambda(z)\langle \pi_Q(x)z, T(\phi) \rangle = 1/1 - \lambda(z) \cdot \langle \pi_Q(x)(1 - z), T(\phi) \rangle$ for all $x \in B(\mathfrak{H})$. Take a directed set $\{\pi_Q(x_\alpha)\}$ such that $\| x_\alpha \| \leqq 1$ and $\pi_Q(x_\alpha) \to z$ (strongly) in N, then

$$\langle \pi_Q(x_\alpha)z, T(\phi) \rangle \to \langle z, T(\phi) \rangle = \langle T^*(z), \phi \rangle = \lambda(z) \neq 0.$$

On the other hand, $\langle \pi_Q(x_\alpha)(1 - z), T(\phi) \rangle \to \langle z(1 - z), T(\phi) \rangle = 0$. This is a contradiction; hence $\lambda(z) = 0$ or 1.

Suppose that $\lambda(z) = 0$, then $\langle z, T(\phi) \rangle = \langle T^*(z), \phi \rangle = \lambda(z) = 0$.

Now consider the *-representation $\{\pi_Q, \mathfrak{H}_Q\}$ and let I_Q be the image of of the identity in $\mathfrak{H}_Q$, then $0 = \langle z, T(\phi) \rangle = (zI_Q, I_Q)$, because $\langle \pi_Q(x), T(\phi) \rangle = \phi_Q(x) = (\pi_Q(x)I_Q, I_Q)$ for $x \in B(\mathfrak{H})$, where $(\ ,\ )$ is the inner product of $\mathfrak{H}_Q$; hence $zI_Q = 0$ and so $\pi_Q(B(\mathfrak{H}))zI_Q = z\pi_Q(B(\mathfrak{H}))I_Q = 0$; hence $z = 0$. Similarly if $\lambda(z) = 1$, we have $z = I$. Therefore N is a factor. This completes the proof.

Now we shall prove the theorem.

Proof of Theorem. Let $[\pi_Q(M)I_Q]$ be the closed subspace of $\mathfrak{H}_Q$ generated by $\pi_Q(M)I_Q$ and E' be the projection of $\mathfrak{H}_Q$ onto $[\pi_Q(M)I_Q]$, then E' belongs to the commutant of $\pi_Q(M)$. $\pi_Q(M)E'$ is weakly closed, because ϕ_Q is normal on M; therefore, $E'\pi_Q(M)'E'$ is a type III-factor, where $\pi_Q(M)'$ is the commutant of $\pi_Q(M)$.

Let F' be the central envelope of E' in $\pi_Q(M)'$ then $F'\pi_Q(M)'F'$ is a type III-factor. Therefore, there is a family of orthogonal, equivalent projections $\{E_\alpha'\}$ in $\pi_Q(M)'$ such that $\sum_\alpha \oplus E_\alpha' = F'$ and $E_\alpha' \sim E'$.

Now we shall show that $\pi_Q(M)F'$ is weakly closed on $\mathfrak{H}_Q$. Let $\{\pi_Q(a_\beta) \mid a_\beta \in M, \| a_\beta \| \leqq 1\}$ be a directed set such that σ-weak $\lim \pi_Q(a_\beta)F' = A$ on $\mathfrak{H}_Q$. Then there is a subdirected set $\{a_{\beta\gamma}\}$ such that $a_{\beta\gamma} \to a$ (σ-weakly) in M for some $a \in M$; therefore, $\pi_Q(a_{\beta\gamma})E' \to \pi_Q(a)E'$ (σ-weakly) on $\mathfrak{H}_Q$. Then,

$$\begin{aligned}\pi_Q(a_{\beta\gamma})E_\alpha' &= \pi_Q(a_{\beta\gamma})V_\alpha'^*E'V_\alpha' \\ &= V_\alpha'^*\pi_Q(a_{\beta\gamma})E'V_\alpha' \to V_\alpha'^*\pi_Q(a)E'V_\alpha' = \pi_Q(a)E_\alpha'\end{aligned}$$

(σ-weakly) on $\mathfrak{H}_Q$, where V_α' is a partial isometry such that $V_\alpha'^*V_\alpha' = E_\alpha'$ and $V_\alpha'V_\alpha'^* = E'$. Therefore,

$$\sum_\alpha \oplus \pi_Q(a_{\beta\gamma})E_\alpha' = \pi_Q(a_{\beta\gamma})F' \to \sum \oplus \pi_Q(a)E_\alpha' = \pi_Q(a)F'$$

(σ-weakly) on $\mathfrak{H}_Q$; hence $\pi_Q(a)F' = A$ and so $\pi_Q(M)F'$ is weakly closed.

Now we shall prove that N is of type III. The proof is quite similar with the case of [5].

Suppose that N is semi-finite, then there is a semi-finite, faithful, normal trace τ on N (cf. [2]). Put $N_0 = \{e \mid \tau(e) < +\infty,\ e \text{ projection in } N\}$. F' belongs to $\pi_Q(M)'' \subset N$; $T^*(F') \neq 0$, because $\langle T^*(F'), \phi\rangle = \langle F', T(\phi)\rangle = (F'I_Q, I_Q)$ and so $T^*(F') = 0$ implies $F'I_Q = 0$, but $I_Q \in E'\mathfrak{H}_Q \subset F'\mathfrak{H}_Q$. Therefore, there is a non-zero $e_0 \in N_0$ such that $e_0 \leqq F'$ and $T^*(e_0) \neq 0$, and so there is a non-zero projection p in M such that $\lambda p \leqq T^*(e_0)$ for some positive number λ.

Suppose that a directed set (a_α) $(\| a_\alpha \| \leqq 1, a_\alpha \in pMp)$ is strongly convergent to 0 in M, then $\pi_Q(a_\alpha)F'$ is strongly convergent to 0 on $\mathfrak{H}_Q$, because $\pi_Q(M)F'$ is weakly closed and so the mapping $a \to \pi_Q(a)F'$ $(a \in M)$ is a W^*-representation; hence $\pi_Q(a_\alpha)e_0$ converges to 0 strongly on $\mathfrak{H}_Q$; by the finiteness of e_0, $\{e_0\pi_Q(a_\alpha)^*\}$ is strongly convergent to 0 (cf. [5], [6]).

Then,

$$\begin{aligned}&T^*((e_0\pi_Q(a_\alpha)^*)^*(e_0\pi_Q(a_\alpha)^*)) \\ &\qquad \geqq T^*(e_0\pi_Q(a_\alpha)^*)^*T^*(e_0\pi_Q(a_\alpha)^*) \to 0 \quad (\sigma\text{-weakly})\end{aligned}$$

in M; hence $\{T^*(e_0\pi_Q(a_\alpha)^*)\}$ converges strongly to 0 in M. Therefore,

$$\begin{aligned}&\{pT^*(e_0)p + (I-p)\}^{-1}pT^*(e_0\pi_Q(a_\alpha)^*) \\ &\qquad = \{pT^*(e_0)p + (I-p)\}^{-1}pT^*(e_0)a_\alpha^* \\ &\qquad = a_\alpha^* \to 0 \ (\text{strongly}) \text{ in } M.\end{aligned}$$

Hence, the *-operation is strongly continuous on bounded spheres of pMp, but pMp is of type III. This is a contradiction (cf. [5], [6]). This completes the proof of the theorem.

Remark. The discussions which are used in the proof of the theorem are applicable to more general situation. Using those discussions and the result of Glimm concerning separable type I C^*-algebras, we can show that every C^*-algebra containing a non-type I separable C^*-subalgebra has a type III-factor *-representation. This will be proved in the next paper.

UNIVERSITY OF PENNSYLVANIA.

REFERENCES.

[1] J. W. Calkin, "Two sided ideals and congruences in the ring of bounded operators in Hilbert space," *Annals of Mathematics*, vol. 42 (1941), pp. 839-873.

[2] J. Dixmier, *Les algèbras d'operateurs dans l'espace hilbertien*, Gauthier-Villars, Paris, 1957.

[3] A. Grothendieck, "Produit tensoriels topologiques et espaces nucléaires," *Memoirs of the American Mathematical Society*, 16.

[4] S. Sakai, "On the σ-weak topology of W^*-algebras," *Proceedings of the Japanese Academy*, vol. 32 (1956), pp. 329-332.

[5] ———, "On topological properties of W^*-algebras," *Proceedings of the Japanese Academy*, vol. 33 (1957), pp. 439-444.

[6] ———, *The theory of W^*-algebras*, Lecture notes, Yale University, 1962.

[7] J. Schwartz, "Two finite, non-hyperfinite, non-isomorphic factors," *Communications in Pure and Applied Mathematics*, vol. 16 (1963), pp. 19-26.

[8] ———, "Non-isomorphism of a pair of factors of type III," *Communications in Pure and Applied Mathematics*, vol. 16 (1963), pp. 111-120.

[9] Z. Takeda, "Conjugate spaces of operator algebras," *Proceedings of the Japanese Academy*, vol. 30 (1954), pp. 90-95.

[10] J. Tomiyama, "On the projection of norm one in W^*-algebras," *Proceedings of the Japanese Academy*, vol. 33 (1957), pp. 608-612.

EXTREMAL PROBLEMS FOR THE TYPICALLY REAL FUNCTIONS.

By WILLIAM E. KIRWAN.

---

1. Introduction. Let $\mathfrak{P}$ denote the class of functions $p(z)$, analytic in $|z|<1$, which satisfy $\operatorname{Re} p(z)>0$ for $|z|<1$ and $p(0)=1$. In a recent paper [7] Robertson proved the following theorem.

THEOREM 1.1. *Let* $F(u,v)$ *be analytic in the half-plane* $\operatorname{Re} u>0$ *and the v-plane. Then on* $|z|=r<1$, $\max_{p\in\mathfrak{P}} \operatorname{Re}\{F(p(z),p'(z))\}$ *is attained by a function of the form*

$$p(z)=\frac{1+t}{2}\,\frac{1+ze^{i\theta}}{1-ze^{i\theta}}+\frac{1-t}{2}\,\frac{1+ze^{-i\theta}}{1-ze^{-i\theta}}$$

where $-1\leqq t\leqq 1$ *and* $0\leqq\theta\leqq 2\pi$.

This result has been generalized by Sakaguchi [9] and by the author in his doctoral dissertation. The result obtained in the latter is

THEOREM 1.2. *Let* $|c|<1$ *and let* $F(\omega_0,\cdots,\omega_{m+1})$ *be analytic on* $\bigcup_{p\in\mathfrak{P}}(p(c),\cdots,p^{(m)}(c),c)$. *Then* $\max_{p\in\mathfrak{P}}\operatorname{Re}\{F(p(c),\cdots,p^{(m)}(c),c)\}$ *is attained by a function of the form*

$$p(z)=\sum_{k=0}^{m}t_k\,\frac{1+\epsilon_k z}{1-\epsilon_k z}$$

where $\sum_{k=0}^{m}t_k=1$, $t_k\geqq 0$ *and* $|\epsilon_k|=1$.

In this paper we consider a similar problem for the class $\mathfrak{T}$ of typically-real functions. This class of functions was introduced by Rogosinski [8] and consists of the functions $f(z)$, analytic in $|z|<1$, which satisfy $f(0)=0$, $f'(0)=1$ and the condition $\operatorname{Im} f(z)=0$ if and only if $\operatorname{Im} z=0$. We prove the following theorem for $\mathfrak{T}$:

THEOREM 1.3. *Let* $|c|<1$ *and let* $F(\omega_0,\cdots,\omega_{m+1})$ *be analytic on* $\bigcup_{f\in\mathfrak{T}}(f(c),\cdots,f^{(m)}(c),c)$. *Then*

Received September 13, 1965.

$$(1.1) \qquad \max_{f \in \mathcal{T}} (\min) \operatorname{Re}(F(c), \cdots, f^{(m)}(c), c)$$

is attained by a function of the form

$$f(z) = z \sum_{k=0}^{p} t_k (1 - 2s_k z + z^2)^{-1}$$

where $p \leqq m + 1$, $-1 \leqq s_k \leqq 1$, $\sum_{k=0}^{p} t_k = 1$ *and* $t_k > 0$. *It must be assumed that if* $f_0(z) \in \mathcal{T}$ *is an extremal function for* (1.1), *there exists a* k, $0 \leqq k \leqq m$, *such that* $\frac{\partial}{\partial \omega_k} F(\omega_0, \cdots, \omega_{m+1})$ *at* $(f_0(c), \cdots, f_0^{(m)}(c), c)$ *is non-zero.*

The proof of Theorem 1.1 and 1.2 utilized the powerful method of interior variations. Part of the purpose of this paper is to show that in some instances theorems of this type may be proven using more elementary techniques. The proof of Theorem 1.3 is a modification of an argument of Goluzin [1], and requires but two simple variational formulae which are derived in § 2. An adaptation of these formulae to functions in $\mathcal{P}$ will yield an elementary proof of Theorems 1.1 and 1.2. As an application of Theorem 1.3, we prove in § 4 that the radius of univalence and the radius of starlikeness for $\mathcal{T}$ are both equal to $\sqrt{2} - 1$.

It should be noted that variational formulae analogous to those of § 2 have been used by Pfaltzgraff [4] to prove a theorem similar to Theorem 1.2. The reader is also directed to [2] and [10] where related ideas have been used.

2. A dense subclass of $\mathcal{T}$.

Robertson [6] showed that $f(z) \in \mathcal{T}$ if and only if

$$(2.1) \qquad f(z) = \int_{-1}^{1} K(z, s)\, d\alpha(s)$$

where

$$K(z, s) = z(1 - 2sz + z^2)^{-1},$$

$\alpha(s)$ is montonically increasing on $[-1, 1]$ and $\alpha(1) - \alpha(-1) = 1$. If $\alpha(s)$ is a step function with discontinuities at $-1 \leqq s_0 < s_1 < \cdots < s_p \leqq 1$ and jump t_k at s_k, then $t_k > 0$ and $\sum_{k=0}^{p} t_k = 1$. It follows from (2.1) that the corresponding function $f(z)$ may be written

$$(2.2) \qquad f(z) = \sum_{k=0}^{p} t_k K(z, s_k).$$

Let $\mathcal{T}_n$ denote the collection of functions in $\mathcal{T}$ which have the representation (2.2) with $p \leqq n$. It is easy to verify that for each $n \geqq 1$, $\mathcal{F}_n$ is a normal compact family of functions. Further, if $\operatorname{cl} \bigcup_{n=1}^{\infty} \mathcal{F}_n$ denotes the collection of functions which are the limit (uniform on compact subsets of $|z| < 1$) of a sequence of functions from $\bigcup_{n=1}^{\infty} \mathcal{F}_n$, then $\operatorname{cl} \bigcup_{n=1}^{\infty} \mathcal{F}_n = \mathcal{F}$. This follows from (2.1) and the fact that any increasing function $\alpha(s)$ on $[-1, 1]$ is the uniform limit of a sequence of increasing step functions.

If $f(z) \in \mathcal{F}_n$,

$$f(z) = \sum_{k=0}^{p} t_k K(z, s_k) \qquad (p \leqq n),$$

then for each $s_j \neq \pm 1$ and real ϵ with $|\epsilon|$ small,

$$f^*(z) = \sum_{\substack{k=0 \\ k \neq j}}^{p} t_k K(z, s_k) + t_j K(z, s_j + \epsilon)$$

defines a function in $\mathcal{F}_n$.

$$K(z, s_j + \epsilon) = K(z, s_j) + \epsilon K_2(z, s_j) + o(\epsilon)$$

$\left(K_2(z, s) = \dfrac{\partial}{\partial s} K(z, s)\right)$ and therefore

$$f^*(z) = f(z) + \epsilon t_j K_2(z, s_j) + o(\epsilon). \tag{2.3}$$

In the representation (2.2), $t_k > 0$. Thus if ϵ is real, $|\epsilon|$ is small and $0 \leqq j \leqq p - 1$

$$\begin{aligned} f^{**}(z) &= \sum_{\substack{k=0 \\ k \neq j, j+1}}^{p} t_k K(z, s_k) + (t_{j+1} + \epsilon) K(z, s_{j+1}) + (t_j - \epsilon) K(z, s_j) \\ &= f(z) + \epsilon [K(z, s_{j+1}) - K(z, s_j)] \end{aligned} \tag{2.4}$$

defines a function in $\mathcal{F}_n$. (2.3) and (2.4) are variational formulae for $\mathcal{F}_n$ and will be used in the proof of Theorem 1.3. They are a modification of a more general variational formula for $\mathcal{F}$ due to Goluzin [1].

3. Proof of Theorem 1.3. Let $F(\omega_0, \cdots, \omega_{m+1})$ satisfy the hypothesis of Theorem 1.3. For notational purposes denote $\operatorname{Re}\{F(f(c), \cdots, f^{(m)}(c), c)\}$, $f(z) \in \mathcal{F}$, by $J(f)$. Let $f_0(z)$ be an extremal function for $\max_{f \in \mathcal{F}} J(f)$. Since $\operatorname{cl} \bigcup_{1}^{\infty} \mathcal{F}_n = \mathcal{F}$, there exists a sequence of functions $\{f_n(z)\}$, $f_n(z) \in \mathcal{F}_{p_n}$, such that $\lim_n f_n(z) = f_0(z)$ uniformly on compact subsets of $|z| < 1$. For convenience we suppose that p_n has been chosen as the smallest integer p such that $f_n(z) \in \mathfrak{F}_p$. $\frac{\partial}{\partial \omega_k} F(\omega_0, \cdots, \omega_{m+1})$ at $(f_0(c), \cdots, f_0^{(m)}(c), c)$ is non-zero for some k, $0 \leqq k \leqq m$, and therefore $\frac{\partial}{\partial \omega_k} F(\omega_0, \cdots, \omega_{m+1})$ at $(f_n(c), \cdots, f_n^{(m)}(c), c)$ is non-zero for some k, $0 \leqq k \leqq m$, and all n sufficiently large. By discarding a finite number of terms of the sequence $\{f_n(z)\}$ it may be assumed that $\frac{\partial}{\partial \omega_k} F(\omega_0, \cdots, \omega_{m+1})$ at $(f_n(c), \cdots, f_n^{(m)}(c), c)$ is non-zero for all n. It may also be assumed that

$$\max_{f \in \mathfrak{F}_{p_n}} J(f) = J(f_n). \tag{3.1}$$

Indeed, if this is not the case select a sequence $\{g_n(z)\}$, $g_n(z) \in \mathfrak{F}_{p_n}$, such that

$$\max_{f \in \mathfrak{F}_{p_n}} J(f) = J(g_n).$$

Since $\mathfrak{F}$ is a compact family of functions, $\{g_n(z)\}$ contains a subsequence $\{g_{n_k}(z)\}$ which converges uniformly on compact subsets of $|z| < 1$. Let $g(z) = \lim_k g_{n_k}(z)$. Then

$$\begin{aligned} J(f_0) &= \lim_k J(f_{n_k}) \\ &\leqq \lim_k J(g_{n_k}) \\ &= J(g). \end{aligned}$$

This inequality implies that $g(z)$ is also an extremal function for $\max_{f \in \mathfrak{F}} J(f)$. Thus if $\{f_n(z)\}$ does not satisfy (3.1) we replace it by the sequence $\{g_{n_k}(z)\}$.

We will show that each $f_n(z) \in \mathfrak{F}_{m+1}$. Since $\mathfrak{F}_{m+1}$ is compact, this will imply $f_0(z) \in \mathfrak{F}_{m+1}$ and the proof will be complete.

Each $f_n(z)$ has the representation

$$f_n(z) = \sum_{k=0}^{p_n} t_{k,n} K(z, s_{k,n})$$

where $t_{k,n} > 0$, $\sum_{k=0}^{p_n} t_{k,n} = 1$ and $-1 \leqq s_{0,n} < \cdots < s_{p_n,n} \leqq 1$. Apply the variational formula (2.3) to $f_n(z)$. For each $s_{j,n} \neq \pm 1$

$$f_n^*(z) = f_n(z) + \epsilon t_j K_2(z, s_{j,n}) + o(\epsilon)$$

defines a function in $\mathfrak{F}_{p_n}$. Applying Taylor's theorem to $J(f_n^*)$ as a function of ϵ we write

$$J(f_n^*) = J(f_n) + \epsilon t_j H_2(c, s_{j,n}) + o(\epsilon)$$

where

$$H(z,s) = \operatorname{Re}\{\sum_{k=0}^{n} F_k(f_n(c), \cdots, f_n^{(m)}(c), c) \frac{\partial^k}{\partial z^k} K(z,s)\} \tag{3.2}$$

and

$$H_2(z,s) = \frac{\partial}{\partial s} H(z,s).$$

$f_n(z)$ is an extremal function for $\max_{f \in \mathfrak{F}_{p_n}} J(f)$ and consequently

$$\epsilon H_2(c, s_{j,n}) + o(\epsilon) \leqq 0.$$

Since ϵ can assume both positive and negative values in this inequality, we conclude $H_2(c, s_{j,n}) = 0$ and that $H_2(c,s)$ vanishes at each $s_{j,n} \neq \pm 1$.

Apply the variational formula (2.4) to $f_n(z)$. For each j, $0 \leqq j \leqq p_n - 1$

$$f_n^{**}(z) = f_n(z) + \epsilon[K(z, s_{j+1,n}) - K(z, s_{j,n})]$$

defines a function in $\mathfrak{F}_{p_n}$. Using Taylor's theorem as before we write

$$J(f_n^{**}) = J(f_n) + \epsilon[H(c, s_{j+1,n}) - H(c, s_{j,n})] + o(\epsilon)$$

where $H(z,s)$ is defined by (3.2). Because of the extremal property of $f_n(z)$

$$\epsilon[H(c, s_{j+1,n}) - H(c, s_{j,n})] + o(\epsilon) \leqq 0$$

for real ϵ with $|\epsilon|$ small. We conclude that $H(c, s_{j+1,n}) = H(c, s_{j,n})$; i.e., $H(c,s)$ has the same value at $s_{j,n}$ and $s_{j+1,n}$ for each j, $0 \leqq j \leqq p_n - 1$. By the mean value theorem $H_2(c,s)$ vanishes on $(s_{j,n}, s_{j+1,n})$. If these zeros are combined with the previously determined zeros of $H_2(c,s)$, it follows that $H_2(c,s)$ has at least $2p_n - 1$ zeros in $[-1,1]$ (we cannot conclude s_0 or s_{p_n} are zeros of $H_2(c,s)$ if they are -1 or 1 respectively). On the other hand $H_2(c,s)$ is a rational function of s. Since $\frac{\partial}{\partial \omega_k} F(\omega_0, \cdots, \omega_{m+1})$ at $(f_n(c), \cdots, f_n^{(m)}(c), c)$ is non-zero for some k, $0 \leqq k \leqq m$, $H_2(c,s)$ does not vanish identically. It is easy to see that $H_2(c,s)$ can have no more than $2m + 2$ zeros. Therefore $2p_n - 1 \leqq 2m + 2$ or $p_n \leqq m + 1$. Thus $f_n(z)$ and hence $f_0(z)$ belong to $\mathfrak{F}_{m+1}$. This completes the proof.

For future reference we observe that either $f_0(z) \in \mathfrak{F}_m$ or in the representation for $f_0(z)$

$$f_0(z) = \sum_{k=0}^{m+1} t_k K(z, s_k)$$

with $s_0 = -1$ or $s_{m+1} = 1$. Indeed, $\lim_n s_{0,n} = s_0$ and $\lim_n s_{p_n,n} = s_{m+1}$. If

$s_0 > -1$ and $s_{m+1} < 1$ then for n sufficiently large $s_{0,n} > -1$ and $s_{p_n,n} < 1$. In this case $s_{0,n}$ and $s_{p_n,n}$ are roots of $H_2(c,s) = 0$ so that this equation has $2p_n + 1$ roots on $[-1,1]$. Thus $2p_n + 1 \leqq 2m + 2$ which implies $p_n \leqq m$ and that $f_n(z)$ and hence $f_0(z)$ belong to $\mathcal{J}_m$.

4. Applications. The following lemma is used in the applications of Theorem 1.3.

LEMMA 3.1. *For fixed* z, $|z| < 1$ *and* $\operatorname{Im} z > 0$, *let*

$$\gamma(t) = (1-z^2)(1-2tz+z^2)^{-1}.$$

On the interval $-1 \leqq t \leqq 1$ $\operatorname{Re}\gamma(t)$ *is either*

a) *strictly increasing,*

b) *strictly decreasing, or*

c) *there exists* t_0, $-1 < t_0 < 1$ *such that* $\operatorname{Re}\gamma(t)$ *is strictly increasing on* $[-1, t_0]$ *and strictly decreasing on* $[t_0, 1]$.

Proof. For fixed z, $|z| < 1$ and $\operatorname{Im} z > 0$, the transformation

$$\gamma(t) = (1-z^2)(1-2tz+z^2)^{-1}$$

maps the real axis onto a circle through the origin and the upper half plane onto the interior of this circle. For fixed t, $-1 \leqq t \leqq 1$, the transformation

$$w = (1-z^2)(1-2tz+z^2)^{-1}$$

maps the unit disc onto the right half plane. These two facts together with elementary geometrical considerations imply the lemma.

Functions in $\mathcal{J}$ of the form

$$f_t(z) = \frac{1-t}{2}K(z,1) + \frac{1+t}{2}K(z,-1) \qquad (-1 \leqq t \leqq 1) \tag{4.0}$$

play an important role in determining the radius of starlikeness of $\mathcal{J}$. The relation between the functions $f_t(z)$ and $\gamma(t)$ is given by the following easily derived formula.

$$\begin{aligned} \frac{zf_t'(z)}{f_t(z)} &= \frac{1+z}{1-z} + \frac{1-z}{1+z} - \frac{1-z^2}{1-2tz+z^2} \\ &= \gamma(1) + \gamma(-1) - \gamma(t). \end{aligned} \tag{4.0}$$

(4.1) will be used in the proof of the next theorem.

THEOREM 4.1. *For fixed* z, $|z|<1$, *let* t_0, $-1\leqq t_0\leqq 1$, *denote the point at which* $\max_{-1\leqq t\leqq 1}\operatorname{Re}\gamma(t)$ *occurs. Then*

$$\min_{f\in\mathcal{T}}\operatorname{Re}\frac{zf'(z)}{f(z)}=\operatorname{Re}\frac{zf_{t_0}'(z)}{f_{t_0}(z)}.$$

Proof. If $f(z)\in\mathcal{T}$, $f(z)$ is real on the real axis. Thus if $\bar z$ denotes the complex conjugate of z,

$$\operatorname{Re}\frac{\bar zf'(\bar z)}{f(\bar z)}=\operatorname{Re}\frac{zf'(z)}{f(z)}.$$

Consequently in the proof of this theorem we need only consider values of θ ($\theta=\arg z$) in the range $0\leqq\theta\leqq\pi$. Moreover, for real values of z it is known [6] that if $f(z)\in\mathcal{T}$,

$$\frac{1-|z|}{1+|z|}\leqq\operatorname{Re}\frac{zf'(z)}{f(z)}.$$

Since $\dfrac{1-|z|}{1+|z|}=\operatorname{Re}\dfrac{zf_t'(z)}{f_t(z)}$ with $t=1$ if $0<z$ and $t=-1$ if $z<0$ the theorem follows for real z.

Consider now a fixed z with $\operatorname{Im}z>0$. By Theorem 1.3 $\min_{f\in\mathcal{T}}\operatorname{Re}\{\dfrac{zf'(z)}{f(z)}\}$ is attained by a function in $\mathcal{T}_3$; i.e., a function of the form

(4.2) $f(z)=z[t_1(1-2rz+z^2)^{-1}+t_2(1-2sz+z^2)^{-1}+t_3(1-2tz+z^2)^{-1}]$

where $t_k\geqq 0$, $\sum_{k=1}^{3}t_k=1$ and $-1\leqq r,s,t\leqq 1$. (Note that since t_k may be zero in (4.2) the representation (4.2) includes the functions in $\mathcal{T}_1$ and $\mathcal{T}_2$ as well as $\mathcal{T}_3$.) We will show that for functions $f(z)$ defined by (4.2)

$$\operatorname{Re}\frac{zf'(z)}{f(z)}\geqq\operatorname{Re}\frac{zf_{t_0}'(z)}{f_{t_0}(z)}.\tag{4.3}$$

This will compete the proof.

By the remark at the end of Theorem 1.3 it may be assumed that in (4.2) one of the values r, s and t, say r, is either 1 or -1. Further it is no loss of generality to suppose $s\leqq t$. For functions of the form (4.2) with $r=1$,

$$\begin{aligned}\frac{zf'(z)}{f(z)}=\frac{1+z}{1-z}+\frac{1-z^2}{1-2sz+z^2}+\frac{1-z^2}{1-2tz+z^2}\\ -2\,\frac{1-az+az^3-z^4}{1-2az+2bz^2-2az^3+z^4}\end{aligned}\tag{4.4}$$

where

$$\begin{aligned} a &= t_1(s+t)+t_2(1+t)+t_3(1+s) \\ b &= t_1(1+2st)+t_2(1+2t)+t_3(1+2s) \qquad (4.5) \\ &= 1+2(t_1st+t_2t+t_3s). \end{aligned}$$

Now

$$2\,\frac{1-az+az^3-z^4}{1-2az+2bz^2-2az^3+z^4}=\frac{1-z^2}{1-2c_1z+z^2}+\frac{1-z^2}{1-2c_2z+z^2} \tag{4.6}$$

$$=\gamma(c_1)+\gamma(c_2)$$

where

$$\begin{aligned} c_1 &= \tfrac{1}{2}[a-(a^2-2b+2)^{\frac{1}{2}}] \\ c_2 &= \tfrac{1}{2}[a+(a^2-2b+2)^{\frac{1}{2}}]. \end{aligned} \tag{4.7}$$

A long but straightforward calculation shows that $a^2-2b+2\geqq 0$ and hence c_1 and c_2 are real numbers. More specifically we have

$$s\leqq c_1\leqq t\leqq c_2\leqq 1. \tag{4.8}$$

Once this inequality is established a simple application of Lemma 3.1 will complete the proof.

From (4.7),

$$\tfrac{1}{2}(c_1+c_2)=\tfrac{1}{2}a \quad\text{and}\quad \tfrac{1}{2}(c_1-c_2)=\tfrac{1}{2}(a^2-2b+2)^{\frac{1}{2}}.$$

Thus $c_1\leqq t\leqq c_2$ is equivalent to

$$|\,t-\tfrac{1}{2}a\,|\leqq\tfrac{1}{2}(a^2-2b+2)^{\frac{1}{2}}$$

and since $s\leqq t$, $s\leqq c_1$ is equivalent to

$$|\,s-\tfrac{1}{2}a\,|\geqq\tfrac{1}{2}(a^2-2b+2)^{\frac{1}{2}}.$$

Finally, $c_2\leqq 1$ only if

$$(a^2-2b+2)^{\frac{1}{2}}\leqq 2-a.$$

All three of these inequalities follow easily from the definition of a and b.

Combining (4.4) and (4.6) we may write

$$\frac{zf'(z)}{f(z)}=\gamma(s)-\gamma(c_1)+\gamma(t)-\gamma(c_2)+\gamma(1). \tag{4.9}$$

Because of (4.8) it follows from this equation that if $\operatorname{Re}\gamma(t)$ is strictly increasing on $[-1,1]$

$$\operatorname{Re}\frac{zf'(z)}{f(z)} \geqq \operatorname{Re}\gamma(s) \geqq \operatorname{Re}\gamma(-1) = \operatorname{Re}\frac{zf_t'(z)}{f_t(z)}$$

with $t = 1$. If $\operatorname{Re}\gamma(t)$ is strictly decreasing on $[-1, 1]$

$$\operatorname{Re}\frac{zf'(z)}{f(z)} \geqq \gamma(1) = \operatorname{Re}\frac{zf_t'(z)}{f_t(z)}$$

with $t = -1$. Finally, consider the case that there is a t_0 such that $\operatorname{Re}\gamma(t)$ is strictly increasing on $[-1, t_0]$ and strictly decreasing on $[t_0, 1]$. By (4.8) if $t_0 \leqq t$, $\operatorname{Re}\gamma(t) - \operatorname{Re}\gamma(c_2) \geqq 0$ and hence

$$\begin{aligned}\operatorname{Re}\frac{zf'(z)}{f(z)} &\geqq \operatorname{Re}\gamma(s) - \operatorname{Re}\gamma(c_1) + \operatorname{Re}\gamma(1)\\ &\geqq \operatorname{Re}\gamma(-1) + \operatorname{Re}\gamma(1) - \operatorname{Re}\gamma(t_0)\\ &= \operatorname{Re}\frac{zf_{t_0}'(z)}{f_{t_0}(z)}\end{aligned}$$

(see 4.1). If $t \leqq t_0$, $\operatorname{Re} -\gamma(c_1) + \operatorname{Re}\gamma(t) \geqq 0$ and so

$$\begin{aligned}\operatorname{Re}\frac{zf'(z)}{f(z)} &\geqq \operatorname{Rt}\gamma(s) + \operatorname{Re}\gamma(1) - \operatorname{Re}\gamma(c_2)\\ &\geqq \operatorname{Re}\gamma(-1) + \operatorname{Re}\gamma(1) - \operatorname{Re}\gamma(c_2)\\ &\geqq \operatorname{Re}\gamma(-1) + \operatorname{Re}\gamma(1) - \operatorname{Re}\gamma(t_0)\\ &= \operatorname{Re}.\frac{zf_{t_0}'(z)}{f_{t_0}(z)}.\end{aligned}$$

This establishes (4.3) for functions $f(z)$ given by (4.2) with $r = 1$. If $r = -1$ the above argument carries over with little change and so we consider the proof complete.

COROLLARY 4.1. *For fixed z, $|z| < 1$, let t_1, $-1 \leqq t_1 \leqq 1$, denote the point at which $\min_{-1 \leqq t \leqq 1} \operatorname{Re}\gamma(t)$ occurs. Then*

$$\max_{f \in \mathcal{J}} \operatorname{Re}\frac{zf'(z)}{f(z)} = \operatorname{Re}\frac{zf_{t_1}'(z)}{f_{t_1}(z)}.$$

The radius of starlikeness of a class of functions analytic in the unit disc is defined to be the supremum of all real numbers r such that every function in the class maps $|z| < r$ univalently onto a domain starlike with respect to the origin. It is known [3; page 221] that a function $f(z)$, analytic in $|z| \leqq r$ with $f(0) = 0$ and $f'(0) \neq 0$, maps the disc $|z| < r$ univalently onto a domain starlike with respect to the origin if and only if $\operatorname{Re}\dfrac{zf'(z)}{f(z)} > 0$ on $|z| = r$. With the aid of Theorem 4.1 we can determine the radius of starlikeness for $\mathcal{F}$.

THEOREM 4.2. *The radius of starlikeness of* $\mathcal{F}$ *is* $\sqrt{2} - 1$.

Proof. In view of Theorem 4.1 and the above comments the radius of starlikeness of $\mathcal{F}$ is determined by minimizing as a function of t the roots of smallest modulus of

$$\operatorname{Re}\left\{\frac{zf_t'(z)}{f_t(z)}\right\} = 0. \tag{4.10}$$

$f_t(z)$ is defined by (4.0).

$\operatorname{Re}\dfrac{zf_t'(z)}{f_t(z)} = 0$ with $z = re^{i\theta}$ only if

$$\begin{aligned} &2(1+r^2)[1+r^4-4t\cos\theta(r+r^3)+2(\cos\theta+2t^2)r^2] \\ &\qquad -[(1+r^2)^2-4r^2\cos^2\theta](1+r^2-2tr\cos\theta)=0. \end{aligned} \tag{4.11}$$

The appearance of the term $1+r^2$ in (4.11) suggests the substitution $\sigma = r + 1/r$. The smallest root of (4.11) as a function of $r = |z|$ corresponds to the largest root of

$$\sigma^3 - 6t\cos\theta\sigma^2 + (8t^2 + 12\cos^2\theta - 8)\sigma - 8t\cos^3\theta = 0. \tag{4.12}$$

(4.12) is obtained from (4.11) by dividing (4.11) by r^3 and then making the substitution $\sigma = r + 1/r$.

If $t = 0$, (4.12) reduces to

$$\sigma^3 + (12\cos^2\theta - 8)\sigma = 0.$$

The roots of this equation are $\sigma = 0$ and $\sigma = (8 - 12\cos^2\theta)^{\frac{1}{2}} \leqq 2\sqrt{2}$. (We are of course considering only the real roots of (4.12).) If $\theta = \frac{\pi}{2}$, (4.12) reduces to

$$\sigma^3 + (8t^2 - 8)\sigma = 0.$$

The roots of this equation are $\sigma = 0$ and $\sigma = (8 - 8t^2)^{\frac{1}{2}} \leqq 2\sqrt{2}$. Thus if $t = 0$ or $\theta = \frac{\pi}{2}$ the roots of (4.12) are not larger than $\sigma = 2\sqrt{2}$ and hence the roots of (4.11) are not smaller than $r = \sqrt{2} - 1$. The remainder of the proof consists of showing that the expression

$$\sigma^3 - 6t\cos\theta\sigma^2 + (8t^2 + 12\cos^2\theta - 8)\sigma - 8t\cos^3\theta \tag{4.13}$$

is positive for $\sigma > 2\sqrt{2}$, $-1 \leqq t \leqq 1$ and $0 \leqq \theta \leqq \pi$.

If $t\cos\theta < 0$ each term of (4.13) is positive. Thus we need only consider the case that $t\cos\theta > 0$. By symmetry we may suppose that $0 < t \leqq 1$ and $0 \leqq \theta < \frac{\pi}{2}$. Let

$$F(t,\theta) = \frac{\sigma(\sigma^2 - 8)}{-6t\cos\theta} + \sigma^2 - \left(\frac{4}{3}\frac{t}{\cos\theta} + \frac{2\cos\theta}{t}\right)\sigma + \tfrac{4}{3}\cos^2\theta.$$

(4.13) is positive for $\sigma > 2\sqrt{2}$ if and only if $F(t,\theta) < 0$ for $\sigma > 2\sqrt{2}$.

Let

$$G(t,\theta) = \frac{4}{3}\frac{t}{\cos\theta} + \frac{2\cos\theta}{t} \qquad (0 < t, 0 < \theta < \tfrac{\pi}{2}).$$

$$\frac{\partial G}{\partial t} = \frac{4}{3}\frac{1}{\cos\theta} - \frac{2\cos\theta}{t^2}$$

so that $\frac{\partial G}{\partial t} = 0$ only when $t^2 = \frac{3}{2}\cos^2\theta$ or $t = \frac{1}{2}\sqrt{6}\cos\theta$. For such values of t and θ, $G(t,\theta) = \frac{4}{3}\sqrt{6}$. Since for fixed θ, $0 \leqq \theta < \frac{\pi}{2}$, $\lim_{t\to 0+} G(t,\theta) = +\infty$ and $\lim_{t\to+\infty} G(t,\theta) = +\infty$, $G(t,\theta)$ is decreasing on $(0, \frac{1}{2}\sqrt{6}\cos\theta)$ and increasing on $(\frac{1}{2}\sqrt{6}\cos\theta, +\infty)$. Consequently, $G(t,\theta) \geqq \frac{4}{3}\sqrt{6}$ for $0 < t \leqq 1$ and $\operatorname{arc\,cos}\frac{1}{3}\sqrt{6} \leqq \theta < \frac{\pi}{2}$ and $G(t,\theta) \geqq \left(\frac{4}{3}\frac{1}{\cos\theta} + 2\cos\theta\right)$ for $0 < t \leqq 1$ and $0 \leqq \theta \leqq \operatorname{arc\,cos}\frac{1}{3}\sqrt{6}$. Also

$$\frac{\sigma(\sigma^2 - 8)}{-6\cos\theta} \leqq \frac{\sigma(\sigma^2 - 8)}{-6\cos\theta} \qquad (\sigma > 2\sqrt{2}).$$

With these bounds for $G(t,\theta)$ we have for $\sigma > 2\sqrt{2}$ and $\operatorname{arc\,cos}\frac{1}{3}\sqrt{6} \leqq \theta < \frac{\pi}{2}$

$$\begin{aligned} F(t,\theta) &\leqq \frac{\sigma(\sigma^2 - 8)}{-6t\cos\theta} + \sigma^2 - \tfrac{4}{3}\sqrt{6}\sigma + \tfrac{4}{3}\cos^2\theta \\ &\leqq \frac{\sigma(\sigma^2 - 8)}{-2\sqrt{6}} + \sigma^2 - \tfrac{4}{3}\sqrt{6}\sigma + \frac{8}{9} \\ &= -\frac{1}{2\sqrt{6}}(\sigma - \tfrac{2}{3}\sqrt{6})^3 < 0. \end{aligned}$$

For $0 \leqq \theta < \operatorname{arc\,cos}\frac{1}{3}\sqrt{6}$

$$\begin{aligned} F(t,\theta) &\leqq \frac{\sigma(\sigma^2 - 8)}{-6\cos\theta} + \sigma^2 - \left(\frac{4}{3}\frac{1}{\cos\theta} + 2\cos\theta\right)\sigma + \tfrac{4}{3}\cos^2\theta \\ &= \frac{1}{-6\cos\theta}(\sigma - 2\cos\theta)^3 < 0. \end{aligned}$$

Thus $F(t,\theta) < 0$ when $0 < t \leqq 1$ and $0 \leqq \theta < \frac{\pi}{2}$. This implies no root of

(4.12) is larger than $\sigma = 2\sqrt{2}$ and no root of (4.11) is smaller than $r = \sqrt{2} - 1$. It follows that the radius of starlikeness of $\mathfrak{J}$ is at least as large as $\sqrt{2} - 1$. However if $t = 0$ and $z = i(\sqrt{2} - 1)$ $f_t'(z) = 0$. This implies that the radius of starlikeness of $\mathfrak{J}$ can be no larger than $\sqrt{2} - 1$ and the proof is complete.

COROLLARY 4.2. *If $f(z) \in \mathfrak{J}$ then $f'(z)$ does not vanish in $|z| < \sqrt{2} - 1$. The function $f(z) = \frac{1}{2}[K(z,1) + K(z,-1)]$ shows that $\sqrt{2} - 1$ is the largest number with this property.*

If $G(z)$ is analytic and normalized in $|z| < 1$, real on the real axis, and maps $|z| < 1$ onto a domain convex in the direction of the imaginary axis then $zG'(z)$ is typically real [5].

COROLLARY 4.3. *If $G(z)$ maps $|z| < 1$ onto a domain convex in the direction of the imaginary axis and is real on the real axis then $G(z)$ maps $|z| < \sqrt{2} - 1$ onto a convex domain.*

The following result is an immediate consequence of the preceding theorem.

THEOREM 4.3. *The radius of univalence of $\mathfrak{J}$ is $\sqrt{2} - 1$. That is, if $f(z) \in \mathfrak{J}$, $f(z)$ is univalent in $|z| < \sqrt{2} - 1$ but if $r > \sqrt{2} - 1$ there is a function in $\mathfrak{J}$ which is not univalent in $|z| < r$.*

UNIVERSITY OF MARYLAND,
COLLEGE PARK, MARYLAND

REFERENCES.

[1] G. M. Goluzin, "A variational method in the theory of analytic functions," *American Mathematical Society Translations* (2), vol. 18 (1961), pp. 1-14.

[2] O. Lehto, "On the distortion of conformal mappings with bounded boundary rotation," *Annales Acedemicae Fennicae*, Ser. A I, vol. 124 (1952).

[3] Z. Nehari, *Conformal Mapping*, McGraw-Hill, 1952.

[4] J. A. Pfaltzgraff, "Extremal problems and coefficients regions for analytic functions represented by a Stieltjes integral," *Transactions of the American Mathematical Society*, vol. 115 (1965), pp. 270-282.

[5] M. S. Robertson, "Analytic functions starlike in one direction," *American Journal of Mathematics*, vol. 58 (1936), pp. 465-472.

C. U.

[6] ———, " On the coefficients of typically real functions," *Bulletin of the American Mathematical Society*, vol. 41 (1936), pp. 565-572

[7] ———, " Extremal problems for functions with positive real part and applications," *Transactions of the American Mathematical Society*, vol. 106 (1963), pp. 236-253.

[8] W. Rogosinski, " Uber positive harmonische Entwiklungen und typis-relle Potenzreiben," *Mathematische Zeitschrift*, vol. 35 (1932), pp. 93-121

[9] K. Sakaguchi, "A variational method for functions with positive real part," *Journal of the Mathematical Society of Japan*, vol. 16 (1964), pp. 287-296.

[10] O. Tammi, " On the maximalization of the coefficients of schlict and related functions," *Annales Actdemiae Scientianum Fennicae*, Ser. A I, vol. 114 (1952).